Inferences during Story Comprehension: Cortical Recruitment Affected by Predictability of Events and Working Memory Capacity

Sandra Virtue[1], Todd Parrish[2], and Mark Jung-Beeman[2]

Abstract

■ Although it has been consistently shown that readers generate bridging inferences during story comprehension, little is currently known about the neural substrates involved when people generate inferences and how these substrates shift with factors that facilitate or impede inferences, such as whether inferences are highly predictable or unpredictable. In the current study, functional magnetic resonance imaging (fMRI) signal increased for highly predictable inferences (relative to events that were previously explicitly stated) bilaterally in both the superior temporal gyri and the inferior frontal gyri. Interestingly, high working memory capacity comprehenders, who are most likely to generate inferences during story comprehension, showed greater signal increases than did low working memory capacity comprehenders in the right superior temporal gyrus and right inferior frontal gyrus. When comprehenders needed to draw unpredictable inferences in a story, fMRI signal increased relative to explicitly stated events in the left inferior gyrus and in the middle frontal gyrus, irrespective of working memory capacity. These results suggest that the predictability of a text (i.e., the causal constraint) and the working memory capacity of the comprehender influence the different neural substrates involved during the generation of bridging inferences. ■

INTRODUCTION

To successfully comprehend a story, individuals often need to make connections between information that is currently stated in a text with information that was previously mentioned in a text. Consider the following example: ''From the gate, Walter could see his grandmother coming towards him. After she walked away, he knew that his cheeks would be sore for days.'' To understand the last sentence, individuals must generate the inference that Walter's grandmother pinched his cheeks. When comprehenders encounter such a gap in understanding (i.e., a coherence break), they need to connect information about the causes and consequences of text events. In other words, comprehenders need to generate bridging inferences (also known as backward inferences) to establish a coherent representation of a text (e.g., Graesser, Singer, & Trabasso, 1994; van den Broek, 1990). Although behavioral research has revealed various factors that influence the likelihood of people drawing bridging inferences (Linderholm & van den Broek, 2002; Just & Carpenter, 1992; McKoon & Ratcliff, 1992; Singer, Andrusiak, Reisdorf, & Black, 1992), relatively little research has examined the component processes or neural substrates involved in generating such inferences. The goal of this program of research is to use functional magnetic resonance imaging (fMRI) to ultimately provide insight into the specific brain areas and cognitive mechanisms that are involved during bridging inference generation.

One factor that influences how easily comprehenders generate bridging inferences is the causal structure of a text. It is well known that people draw inferences when necessary to maintain story coherence (e.g., Graesser et al., 1994; van den Broek, 1990), as in the example above. In addition, the more the text constrains or limits the number of possible inferences readers can generate during story comprehension, the more likely people are to draw these types of inferences (van den Broek, 1990). For example, when reading ''From the gate, Walter could see his grandmother coming towards him,'' individuals are more likely to generate the inference ''pinch'' after reading ''Walter could see her hands reaching out for his cheeks'' than after ''Walter could see her hands reaching out for him.'' Among other effects, the level of causal constraint in a text appears to differently affect how the right and left hemispheres contribute to bridging inferences. Specifically, weakly constrained text shows greater semantic priming (i.e., speeded responses to related words, indicating greater ''semantic activation'' of inference-related information) for bridging inferences in the right hemisphere than in the left hemisphere (Virtue, van den Broek, & Linderholm, 2006). Manipulating causal constraint then could reveal specific neural components involved in bridging inference generation.

[1]DePaul University, Chicago, IL, [2]Northwestern University, Evanston, IL

Therefore, in the current study, we investigated the neural activity that occurs during the generation of bridging inferences under two levels of causal constraint: when inferences are highly predictable in a story versus when inferences are unpredictable in a story.

Another factor that has been shown to influence how and when people draw bridging inferences is the working memory capacity of the comprehender. Previous research has shown that high working memory comprehenders are more likely than low working memory comprehenders to generate inferences (Linderholm, 2002; Estevez & Calvo, 2000; St. George, Mannes, & Hoffman, 1997). In addition, a comprehender's working memory capacity influences how inferences are generated under different levels of causal constraint (e.g., Virtue, Haberman, Clancy, Parrish, & Jung Beeman, 2006; Virtue, van den Broek, et al., 2006; Linderholm, 2002). Specifically, high working memory capacity readers show more semantic priming when given strongly constrained text than when given weakly constrained text, in both hemispheres. In contrast, low working memory capacity readers show this same pattern in the left hemisphere (i.e., when words are presented to the right visual field), but equivalent semantic priming of inference-related information for both levels of constraint in the right hemisphere (Virtue, van den Broek, et al., 2006). Therefore, in the current study, we examined the relation between working memory capacity, causal constraint, and brain activity associated with drawing bridging inferences.

Although behavioral research has shown that the causal constraint of a text and the working memory capacity of a reader both separately influence bridging inference generation, relatively little research has examined the neural substrates of these effects. Recent fMRI studies have provided some insights into the brain areas that may be involved when people generate bridging inferences. For example, when reading sentence pairs that vary in degree of causal relatedness, different brain areas show distinct patterns of activity (i.e., a dissociation analogous to one previously reported in behavioral studies of reading). Specifically, people read increasingly slowly as the degree of semantic relatedness between the sentences decreases (Myers, Shinjo, & Duffy, 1987). These monotonically increasing reading times have been attributed to the increasing difficulty in establishing connections between sentences as they decrease in relatedness. In contrast, people remember moderately related sentence pairs the best, and remember highly related and unrelated sentence pairs the least well. The improved memory for moderately related pairs has been attributed to the successful generation of causal inferences, which are not necessary for highly related sentences (Myers et al., 1987). Analogously, findings from fMRI studies reveal that activity in the dorsal lateral prefrontal cortex (especially in the left hemisphere) increases monotonically as causal relatedness decreases, whereas activity in the right hemisphere (especially in the right temporal lobe) peaks for moderately related sentence pairs, which are the type of sentence pairs known to elicit causal inferences (Mason & Just, 2004). Consistent with this idea, other research has found superior temporal gyrus (STG) and inferior frontal gyrus (IFG) activity at different times during inference generation (Virtue, Haberman, et al., 2006; Jung-Beeman, 2005) and during semantic integration in other types of tasks (Ferstl, Rinck, & von Cramon, 2005; Jung-Beeman et al., 2004). For example, fMRI signal increases in the STG during early stages of inference generation (i.e., when the verb either infers or explicitly states an inference), whereas fMRI signal increases in the IFG during later stages of inference generation (i.e., when a coherence break occurs in a text) (Virtue, Haberman, et al., 2006). This suggests that the STG may be involved in integrating semantic information, whereas the IFG is involved in selecting semantic information during inference generation (e.g., the Bilateral Activation Integration and Selection [BAIS] model; Jung-Beeman, 2005). Furthermore, the increased neural activity in the IFG for inference events relative to explicit events occurs more strongly in high working memory capacity comprehenders—who are most likely to complete inferences—than in low working memory capacity comprehenders (Virtue, Haberman, et al., 2006).

Taken together, these behavioral and neuroimaging findings lead to several hypotheses about how the brain processes inferences that are highly predictable versus those that are unpredictable. It is possible that when comprehenders encounter text that provides hints about a future inference in a text (i.e., a highly predictable inference), they likely carry out a variety of cognitive processes. Therefore, with highly predictable inferences, by the time comprehenders encounter a coherence break, they already have access to information that will direct them to the appropriate inference. They should be able to integrate and select the necessary semantic information when they reach the coherence break and need to draw the bridging inference. Areas within the STG have been implicated during the semantic integration of inferential information (Virtue, Haberman, et al., 2006; Jung-Beeman, 2005; Mason & Just, 2004; St. George, Kutas, Martinez, & Sereno, 1999), whereas areas within the IFG have been implicated during semantic selection (Barch, Braver, Sabb, & Noll, 2000; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997). For example, activity in both the STG and the IFG are evident when individuals need to recruit additional resources to complete a task, such as generating unusual uses for a noun (Seger, Desmond, Glover, & Gabrieli, 2000). Thus, we predict that when comprehenders hear stories that strongly suggest an event (compared to when stories explicitly state an event), they will show increased neural activity bilaterally in the STG and the IFG when they need to generate a bridging inference (i.e., at the coherence break).

In contrast, with unpredictable inferences, no prior information directs comprehenders to the necessary inference when they encounter a coherence break. Therefore, comprehenders may experience an element of surprise when they reach the coherence break and need to generate the inference. Under these less constrained conditions, comprehenders need to search for potential relations and select the appropriate inference to resolve the unexpected break in coherence. Areas within the IFG have been suggested to be important for directed semantic retrieval and selection. For example, activity in the IFG increases as participants choose between several alternative answers (Wagner, Paré-Blagoev, Clark, & Poldrack, 2001; Barch et al., 2000; Thompson-Schill, et al., 1997) or generate specific uses for a noun (Seger et al., 2000; Posner, Petersen, Fox, & Raichle, 1988). Thus, we predict that when comprehenders need to generate a bridging inference (i.e., at the coherence break), they will show greater neural activity in the left IFG when presented with unpredictable inferences than when presented with text that explicitly states an event during story comprehension.

In summary, we predict that comprehenders will show greater neural activity when listening to stories that imply events than when listening to stories that explicitly state events. We further hypothesize that different cognitive processes, and therefore different neural substrates, will be engaged when comprehenders generate bridging inferences under highly predictable versus unpredictable text conditions. In fact, it is possible that more processes will be engaged (or will be engaged more immediately and consistently) for the highly predictable events than for the unpredictable events because information necessary to complete the inferences is more accessible to comprehenders when the inferences are relatively predictable.

Specific predictions can also be made with regard to how working memory capacity and causal constraint interact during the generation of bridging inferences. Because high working memory comprehenders have additional capacity to store and access semantic information, they are likely able to simultaneously carry out several cognitive processes during the comprehension of text (e.g., semantic integration and selection), which may not be possible by low working memory comprehenders. These additional processes are mostly likely to be evident for highly predictable inferences, as comprehenders are most likely to generate inferences under highly predictable text conditions (van den Broek, 1990). Therefore, it is possible that additional cognitive processes are carried out by high working memory comprehenders during inference generation when the text is highly predictable. Importantly, these unique cognitive processes carried out by high working memory comprehenders should be evident in additional neural activity during bridging inference generation. This hypothesis is consistent with behavioral findings showing that high working memory readers show greater facilitation for inference-related target words in the right hemisphere when the text is strongly constrained than when it is weakly constrained text, whereas low working memory readers show equal facilitation for inference-related targets in the right hemisphere for these two levels of constraint (Virtue, van den Broek, et al., 2006). Thus, high working memory individuals seem to have unique semantic activation in the right hemisphere that assists them during the comprehension of text. These findings also suggest that high working memory comprehenders make use of this increased right hemisphere activation primarily during the processing of strongly constrained text. Thus, we predict that when high working memory capacity comprehenders need to generate a bridging inference (i.e., at the coherence break), they will show greater neural activity in the right hemisphere (particularly in the STG and the IFG) when presented with highly predictable inferences than when presented with explicitly stated inferences during story comprehension. It is less clear how high and low working memory comprehenders might differ when hearing stories implying less predictable events because, by design, these events should not be inferred prior to the coherence break.

In the current study, we investigated neural activity as comprehenders generate bridging inferences that are highly predictable and unpredictable in a story. We use an event-related fMRI design to measure neural activity at the point at which comprehenders need to draw a bridging inference to establish causal coherence in a text (i.e., at the coherence break). In addition, we also examined this neural activity separately for high working memory capacity comprehenders, or those individuals who are most likely to show evidence of bridging inference generation, to investigate how working memory capacity and the predictability of a text (i.e., the causal constraint) interact during story comprehension.

METHODS

Participants

Twenty-six participants (12 men and 14 women, age range = 19–34 years) were native English speakers, right-handed (M = 0.78, range = 0.2–1.0), as assessed by a five-item preference questionnaire (with a range of +1.0 to −1.0) (Bryden, 1982), and without significant history of neurological disorders. Seven participants were replaced prior to analyses: two participants due to poor performance on the comprehension questions, four participants due to scanner error, and one participant whose scan indicated uncorrectable motion. The Northwestern University Institutional Review Board approved all experimental procedures, and written informed consent was obtained from each participant.

Stories and Task during Scanning

During fMRI scanning, participants listened to 20 stories over headphones via an MRA audio system. The 20 stories were divided into three scanning runs (or short sessions) of six or seven stories. Each story was presented in its entirety (approximately 2 to 3 min), and there was no overt task during story comprehension; participants were instructed to listen to the stories and to be prepared to answer multiple-choice questions about the stories (not referring to the inferred or explicit events). These questions occurred following each of the three scanning runs. Each story contained four episodes, with a central causal event in each episode. For each of the four episodes, we created three versions: one in which the event was unpredictable (i.e., there were no hints provided about the future bridging inference), one in which the event was highly predictable (i.e., there were hints provided in the text about the future bridging inference), and one in which the event was explicitly stated (see Table 1). These variations in the text occurred in the sentence prior to the "coherence break," the point at which the inference must be generated (in the two inference conditions). Data from a pilot study in Virtue, van den Broek, et al. (2006) confirmed that participants generated these inferences more frequently following the highly predictable text than following the unpredictable text for all inference events presented in the 20 stories. In addition, both predictable and unpredictable texts led pilot participants to generate the target inference events more often than any other events and reliably more often than zero [$t(47) = 16.40$, $SE = 0.59$, $p < .05$; $t(47) = 4.39$, $SE = 0.35$, $p < .05$, respectively]. We created three material sets from these texts. In each of the three material sets, participants heard 26 or 27 episodes sequentially for each of three conditions: unpredictable inference events, highly predictable inference events, and explicit events. Thus, each participant heard 20 stories containing a mix of highly predictable, unpredictable, and explicit events. Participants heard each of the four episodes of a story in only one of the three conditions. Within each of the four episodes per story, we examined fMRI signal at the coherence break (i.e., the moment the bridging inference needs to be generated to establish coherence in a text). Note that the texts themselves are identical at this time point, the variations having occurred 3 to 9 sec earlier. The critical events (i.e., coherence break for each episode) were well separated by more than 15 sec.

Reading Span Task

Outside of the scanner, all participants performed a reading span task to assess working memory capacity. In this task, participants read aloud a set of sentences that included between two and six sentences. After each set, participants were asked to say aloud the last word in each sentence of the set. After recalling the last word in each sentence of the set, participants were then asked to recall two missing words from one of the sentences in the set they had just seen previously (Singer & Ritchot, 1996). The total number of final words recalled in the reading span task was used to distinguish low from high working memory capacity individuals (Linderholm & van den Broek, 2002). Using the upper and lower half of the distribution of final words recalled to distinguish these two groups [$t(24) = 7.65$, $p < .05$], 13 low working memory capacity readers recalled 27 to 32 total words ($M = 29.86$, $SE = 0.51$) and 13 high working memory capacity individuals recalled 33 to 41 total words ($M = 36.77$, $SE = 0.73$) out of 44 final words presented to participants.

fMRI Data Acquisition

Imaging was performed at Northwestern University's Center for Advanced MRI using a 3-Tesla Siemens Trio scanner and a standard transmit/receive head coil. Head motion was restricted with plastic calipers built into the coil, and a vacuum pillow. Anatomical high-resolution T1-weighted images were acquired in the axial plane parallel to the AC–PC plane at the end of every session. Functional images were acquired in the same axial plane as the anatomical images using a gradient-echo, echo-planar sequence (TR = 2 sec allowing 38 3 mm slices, TE = 20 msec, matrix size 64×64 in 220 mm field of view). The imaging sequence was optimized for detection of the blood oxygenation level-dependent effect (Binder et al., 1999) including local shimming and 8 sec of scanning prior to data collection to allow the MR signal to reach equilibrium. Participants listened to stories in three runs (run 1 = 16 min 40 sec; run 2 = 16 min 40 sec; run 3 = 14 min 17 sec); across three runs a total of 1417 whole-brain volumes were acquired. The total functional scan session took approximately 46 min 97 sec, with an additional 7 min to collect the anatomic scan.

fMRI Data Analysis

Images were coregistered through time using a three-dimensional registration algorithm (Cox, 1996). Echo-planar imaging volumes were spatially smoothed by using a 7.5-mm full-width half-maximum Gaussian kernel to accommodate residual anatomical differences across participants. Within each run, voxels were eliminated if the signal magnitude changed more than 30% between time points (TR = 2 sec) or if the mean signal level was below a threshold defined by the inherent noise in the data acquisition. Finally, all of the runs were transformed (Collins, Neelin, Peters, & Evans, 1994) to conform to a standard stereotaxic atlas (Talairach & Tournoux, 1988), with a final isometric voxel size of 2.5 mm^3.

The data were analyzed with event-related analyses, using general linear model (GLM) analysis (D. Ward, "Deconvolution Analysis of fMRI Time Series Data," http://afni.nimh.nih.gov/afni) that extracted average estimated

Table 1. Sample Story When Bridging Interference Events are Highly Predictable, Unpredictable, or Explicitly Stated

Grandma Johnson was flying in from Florida for a visit.

Her plane had just landed at the terminal and her grandson Walter was looking forward to seeing her.

From the gate, Walter could see his grandmother coming towards him.

<u>Highly Predictable Inference Condition</u>

As Grandma Johnson approached, Walter could see her hands reaching out for his cheeks.

<u>Unpredictable Inference Condition</u>

As Grandma Johnson approached, Walter could see her hands reaching out for him.

<u>Explicit Condition</u>

As Grandma Johnson approached, Walter could tell she was going to pinch his cheeks.

Even though he loved his grandmother and was glad to see her, Walter knew that for days afterward his cheeks would be sore.*

When they got home, Walter brought out his toy dump truck to show his grandmother.

He liked to fill the truck with dirt from the plants and then dump the dirt out the window.

As Walter was showing his grandmother how to dump the dirt onto the ground below, the truck slipped out of his hands.

<u>Highly Predictable Inference Condition</u>

His fragile wooden truck fell out of the third floor window.

<u>Unpredictable Inference Condition</u>

His favorite wooden truck fell out of the kitchen window.

<u>Explicit Condition</u>

His wooden toy truck fell out of the window and broke.

Much to Grandma Johnson's dismay, she looked out the window and saw that Walter's toy dump truck was on the ground, scattered in many different pieces.*

When Grandma Johnson went outside to pick up the pieces of the truck, Walter decided to hunt around in the kitchen closet.

He found several jars on the first shelf in the pantry that were within his reach.

One of them looked especially interesting, so he removed the top.

<u>Highly Predictable Inference Condition</u>

As Walter stuck his nose into the jar, he found that it was full of spicy pepper.

<u>Unpredictable Inference Condition</u>

As Walter stuck his nose into the jar, he found that it was full of sweet cinnamon.

Table 1. (*continued*)

<u>Explicit Condition</u>

As Walter stuck his nose into the jar, he found that the pepper made him sneeze.

Next his grandmother came back inside after hearing some noise in the kitchen, handed Walter a tissue and said, "Bless you!"*

Grandma Johnson spent a week with Walter and his family.

At the end of her visit, they all took her to the airport to send her off.

Walter's family knew how attached Walter was to his grandmother.

<u>Highly Predictable Inference Condition</u>

When it was time to go, Walter's mother handed him a handkerchief.

<u>Unpredictable Inference Condition</u>

When it was time to go, Walter's mother walked over and picked him up.

<u>Explicit Condition</u>

When it was time to go, Walter's mother picked him up because he was crying.

Although his family tried their best to comfort him, Walter covered his face with his hands and realized his cheeks were all wet.*

Walter and his family stayed at the airport to watch Grandma Johnson's plane take off and then went back home.

Note the coherence break* (when comprehenders must generate the inference for successful comprehension) is labeled for each of the four bridging inference events in this story.

responses to each trial type, correcting for linear drift and removing signal changes correlated with head motion. For each participant, average signal responses were estimated for each of three trial types. Within each participant, estimated signal for each inference type was contrasted with signal for explicit events (i.e., highly predictable inferences minus explicit events; and unpredictable inferences minus explicit events). The critical signal was from three TRs, reflecting 4 to 10 sec following the coherence break. Second-order analyses examined whether these contrasts were reliable across participants. Contrast effects across subjects were considered reliable if a cluster larger than 500 mm^3 occurred in which all voxels showed an effect at $p < .001$. The t values to meet this criterion were $t(1, 25) = 3.72$ for contrasts across all participants, $t(1, 12) = 4.29$ for either high working memory or low working memory subgroups alone, and $t(1, 12) = 3.73$ for the contrast between high working memory and low working memory subgroups. (For completeness, Table 2 includes smaller

Table 2. All Areas Larger than 400 mm^3 at $p < .001$ Showing Stronger fMRI Signal for Text Events that Highly Predicted or Did Not Predict the Bridging Inference Compared to Text Events that Explicitly Stated the Bridging Inference at the Coherence Break

Critical Event	*Gyrus/Structure*	*Brodmann's Area*	*Volume*	*Center Coordinates*			*Mean % Signal*	*Maximum % Signal*	*Mean t*	*Maximum t*	*Maximum t Coordinates*		
				x	*y*	*z*					*x*	*y*	*z*
All Participants													
Predictable > Explicit	left middle temporal gyrus	21, 39	3844	−55	−53	12	0.06	0.09	4.2	5.3	−59	−56	18
	right superior temporal gyrus	22, 39	3094	51	−54	16	0.06	0.09	4.1	5.4	51	−51	13
	left inferior frontal gyrus	9	2359	−46	10	27	0.06	0.09	4.1	5	−34	1.5	35.5
	left inferior frontal gyrus	47	938	−48	29	1	0.07	0.10	4	4.7	−39	26.5	−2
	left pyramis		734	−17	−77	−27	0.05	0.07	4.2	5.2	−19	−76	−27
	right inferior frontal gyrus	47	688	49	28	−3	0.09	0.12	4.1	5	48.5	31.5	−7
Unpredictable > Explicit	left inferior frontal gyrus	45, 47	641	−52	30	1	0.08	0.10	4.3	5	−54	29	0.5
	left inferior frontal gyrus	44	609	−47	15	19	0.06	0.07	4	4.5	−46.5	16.5	18
	left middle frontal gyrus	9	469	−38	9	33	0.05	0.07	4.2	5.1	−36.5	9	33
Explicit > Unpredictable	left middle frontal gyrus	9	891	−24	28	30	0.04	0.05	4.3	−5.3	−24	24	30.5
High Working Memory Participants													
Predictable > Explicit	right superior temporal gyrus	22	3922	51	−57	19	0.09	0.14	5.4	8.6	51	−58.5	20.5
	right inferior frontal gyrus	47, 45	1016	48	28	−2	0.10	0.15	4.9	6.2	46	24	−4.5
	right inferior frontal gyrus	9	438	53	8	27	0.09	0.11	4.6	5.3	56	6.5	28
Low Working Memory Participants													
Explicit > Predictable	right parahippocampal gyrus	19	453	29	−50	−3	0.04	0.05	5.1	−7.2	28.5	−51	−4.5
High–Low Working Memory Participants													
Predictable > Explicit	right middle temporal gyrus	39	1109	47	−65	20	0.11	0.13	4.1	5	53.5	−56	15.5

All clusters are shown in descending order of average percent signal change. Location of cluster centers and peak *t* values are shown in Talairach coordinates.

clusters that did not quite meet the cluster size threshold so as not to claim that there were null effects in such areas.)

RESULTS

Comprehension Task Performance

To encourage comprehension, after each imaging run we asked participants a total of 40 multiple-choice questions about the stories (i.e., 14 questions in run 1; 14 questions in run 2; and 12 questions in run 3). Participants answered, on average, 95% of the comprehension questions correctly (*SE* = 0.26, range = 90–100%).

fMRI Results

When all participants (low working memory and high working memory capacity groups together) comprehended episodes with highly predictable inference events, fMRI signal at the coherence break increased more for implied than for explicit events in two clusters in the left IFG, the right IFG, the left middle temporal gyrus (MTG), and the right STG (see Table 2 and Figure 1). For unpredictable inference events, fMRI signal increased more for implied than for explicit events in two clusters in the left IFG and one cluster in the left middle frontal gyrus (MFG). Other brain areas are highly active throughout story comprehension, which is essentially the baseline to which inference events are compared; some of these could contribute to drawing bridging inferences, without showing increased fMRI signal above this high baseline. However, because no other clusters of voxels reliably differed between explicit events and inference events (of either predictable or unpredictable type), we have no evidence of inference-specific activity in other brain areas.

Previous studies indicate that high working memory comprehenders are more likely to generate inferences than low working memory comprehenders (Linderholm, 2002; Estevez & Calvo, 2000; St. George et al., 1997), and that working memory capacity and causal constraint interact (Virtue, van den Broek, et al., 2006). Thus, we expected to see the best evidence of inference-related processing in the right STG and IFG in high working memory capacity comprehenders. Indeed, in the current study, individuals' working memory scores correlated with the strength of fMRI signal increases in the STG for

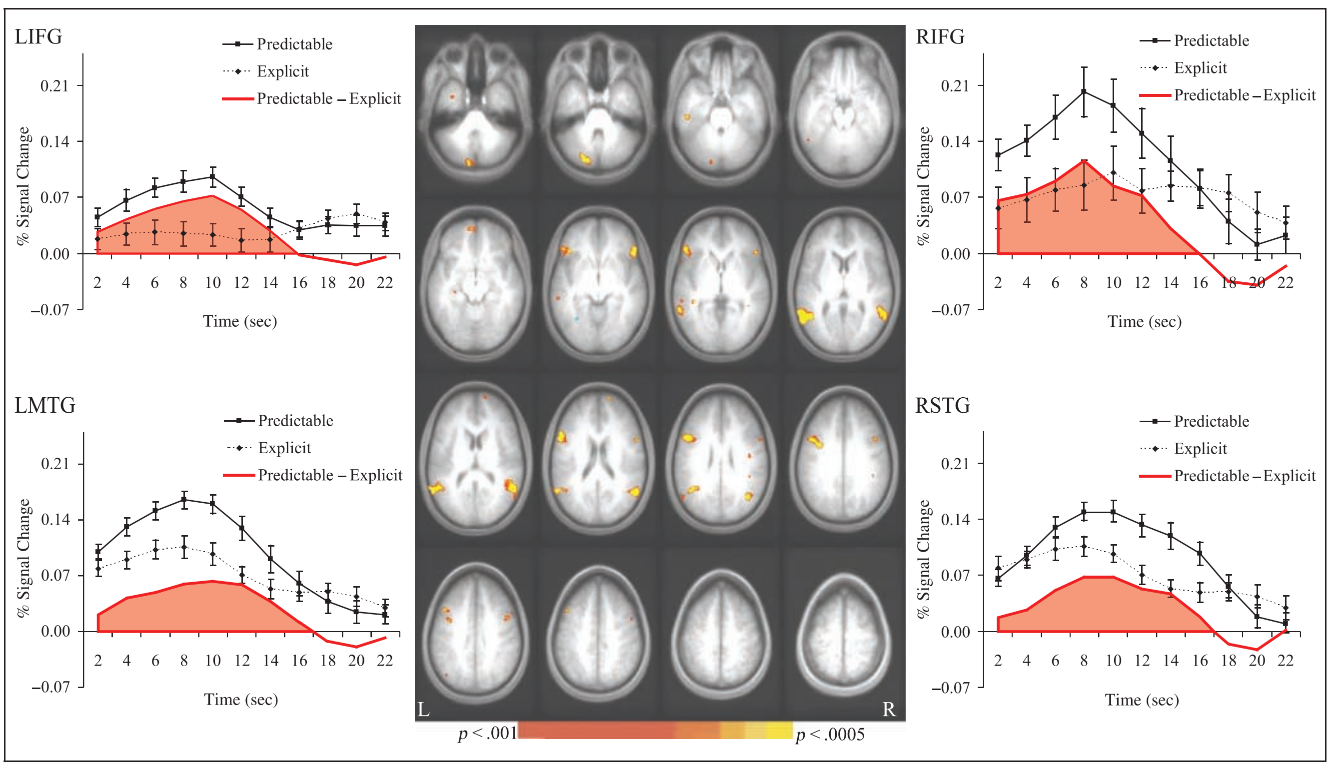

Figure 1. Axial images showing all voxels with stronger signal change (t = 3.72, p < .001) at the coherence break following highly predictable inferences than following explicitly stated events, overlaid on the averaged normalized structural image of all subjects (with left side of brain on left side of images). (See Table 2 for cluster volume size and peak t values.) The graphs show group-average signal change (in % signal change) for the 11 observed time points following highly predictable bridging inference events (solid black line), explicitly stated events (dashed black line), and the difference between the two (filled red line), for the four largest regions (left to right, top to bottom): the left inferior frontal gyrus (LIFG), the right inferior frontal gyrus (RIFG), the left middle temporal gyrus (LMTG), and the right superior temporal gyrus (RSTG). Error bars represent the standard error of the mean signal change at each time point. Note: Graphs are shown for only one of the two significant clusters in the LIFG.

highly predictable inference events relative to explicitly stated events at the coherence break ($r = .51$, $p < .05$; one-tailed). Therefore, we separately analyzed the fMRI data of the high working memory group to identify clusters showing stronger fMRI signal for highly predictable inference events than for explicitly stated events. We also directly contrasted fMRI signal change for highly predictable inference events relative to explicitly stated events in the high working memory group to that of the low working memory group. Although our criterion is set at clusters above 500 mm^3, for completeness and to show clusters just below this threshold, all clusters for these comparisons at the coherence break over 400 mm^3 at $p < .001$ are reported in Table 2.

For the high working memory capacity comprehenders, greater fMRI signal was found for highly predictable inference events than explicitly stated events in several areas in the right hemisphere: the right STG and two clusters in the right IFG (see Table 2). For highly predictable inference events in the low working memory capacity group, one cluster approached significance but was not in any predicted area. The direct contrast between high and low working memory groups for highly predictable inference events compared to explicitly stated events revealed evidence of greater fMRI signal for high working memory participants in the right MTG (see Table 2 and Figure 2). For unpredictable inference events compared to explicitly stated events, no reliable clusters were found in either the high working memory group or in the low working memory group, nor were differences found across the two groups.

DISCUSSION

As expected, the predictability of inferable events influenced the neural activity involved when people generate bridging inferences to maintain coherence. When comprehenders encountered a coherence break under highly predictable conditions, neural activity increased bilaterally in the STG and MTG and in the IFG during

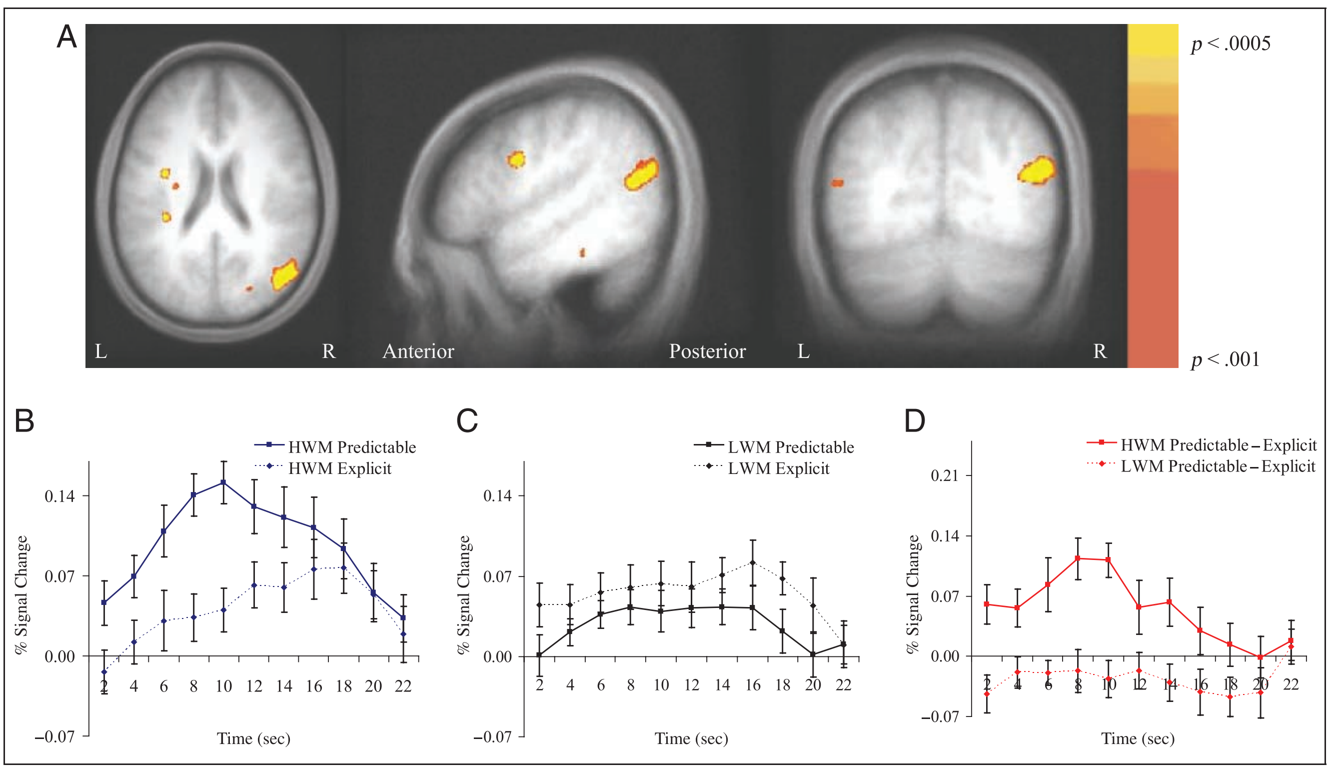

Figure 2. fMRI signal in the right middle temporal gyrus for high working memory (HWM) participants compared to low working memory (LWM) participants for text events that highly predicted the bridging inference. (A) Images showing all voxels with stronger signal change ($t = 3.73$, $p < .001$) at the coherence break following highly predictable inferences than following explicitly stated events for HWM subjects compared to LWM subjects, overlaid on the averaged normalized structural image of 13 HWM subjects and 13 LWM subjects (with left side of the brain on left side of images). (See Table 2 for cluster volume size and peak t values.) (B) Group-average signal change (in % signal change) for the 11 observed time points following highly predictable bridging inference events (blue solid line) and explicitly stated events (blue dashed line) across the entire right middle temporal gyrus for HWM subjects. Error bars represent the standard error of the mean signal change at each time point. (C) Group-average signal change (in % signal change) for the 11 observed time points following highly predictable inference events (green solid line) and explicitly stated events (green dashed line) across the entire right middle temporal gyrus for LWM subjects. Error bars represent the standard error of the mean signal change at each time point. (D) Highly predictable inference event signal minus explicitly stated inference event signal at the coherence break for HWM subjects (red solid line) compared to LWM subjects (red dashed line) across the active region in the right middle temporal gyrus. Error bars represent the standard error of the mean signal change at each time point.

bridging inference generation. In contrast, when comprehenders encountered coherence breaks under less predictable conditions, neural activity increased in the left MFG and IFG. Finally, high working memory participants showed greater neural activity than low working memory participants in the right MTG for highly predictable inferences. In summary, these findings provide additional support and offer new insight into how the predictability of a text (i.e., the causal constraint) and the working memory capacity of a comprehender influence neural activity and bridging inference processing during story comprehension.

The degree of causal constraint produced unique patterns of neural activity during bridging inference generation. From previous behavioral findings, we know that comprehenders are more likely to generate inferences that are strongly constrained by the text than inferences that are weakly constrained by the text (van den Broek, 1990). Consistent with these findings, we found greater neural activity during highly predictable inferences (i.e., bilateral activation in the left MTG, right STG, and left IFG) than during unpredictable inferences (i.e., unilateral activation in the left MFG and left IFG). It is likely that when people read text strongly suggesting an inference, additional resources are used to generate inferences at the coherence break. As proposed earlier, the process of generating bridging inferences that are highly predicted by a text may involve multiple cognitive processes. When comprehenders encounter a potential coherence break following text that strongly suggests an inference event (Virtue, van den Broek, et al., 2006), they take advantage of semantic information they generated when hearing the constraining text to help integrate and select the inference concept. These processes are likely reflected in increased bilateral neural activity during highly predictable inferences.

Interestingly, at the coherence break necessitating unpredictable inferences, neural activity increased predominantly in the left hemisphere, specifically in the left IFG. This increased fMRI signal indicates additional processing as comprehenders attempt to draw unpredictable inferences. Given that the pattern of brain activity differs from that observed for the more predictable inferences (and that it differs from that observed in the most skilled comprehenders), it seems unlikely that this increased neural activity reflects successful inference generation. In fact, prior behavioral studies have shown greater priming for inference-related words presented to the right hemisphere than those presented to the left hemisphere (e.g., Beeman, Bowden, & Gernsbacher, 2000), especially when the inferences are not well constrained (Virtue, van den Broek, et al., 2006). Inference-related priming reflects activation of and/or sensitivity to semantic information necessary to draw an inference, regardless of whether the inference has been completed. In combination, this suggests that when inferences are unpredictable, there may be weak activation of related information, especially in the right hemisphere during semantic processing, and that generating an inference at the coherence break involves contributions from left frontal brain areas. Also, the relatively reduced signal for unpredictable inferences suggests that they are drawn less often, although it is also possible that they are drawn but at varying times after the coherence break (e.g., sometimes quickly and sometimes more slowly) so that they do not elicit consistent fMRI signal.

The working memory capacity of comprehenders was also found to influence neural activity during bridging inference generation following highly predictable text. Whereas low working memory comprehenders showed predominantly left hemisphere activity, high working memory capacity comprehenders showed bilateral neural activity (in the IFG and the STG), including activity in the right MTG that was reliably stronger than in low working memory comprehenders. These findings are consistent with previous studies showing that high working memory readers have greater activation for strongly constrained inferences than for weakly constrained inferences in both hemispheres, whereas low working memory readers only show this pattern in the left hemisphere (Virtue, van den Broek, et al., 2006). Thus, it seems as if high working memory capacity individuals may be carrying out some unique cognitive processes during inference generation, particularly in the right hemisphere. The current study expands upon these findings by providing evidence that the right STG and the right IFG are the specific areas within the right hemisphere which are uniquely processing inferential information in high working memory capacity individuals during story comprehension.

It is important to note that increases in right hemisphere regions were observed in the high working memory participants, for whom story comprehension is generally easier. Indeed, the increased right hemisphere activity in the high working memory participants occurred when they generated the relatively easy inferences (i.e., inferences that were highly predictable in the text). Therefore, these results suggest that these right hemisphere regions may contribute to good story comprehension generally, and good inference generation specifically. Interestingly, this finding is not consistent with the viewpoint that the right hemisphere regions (i.e., homologous to the left hemisphere ''language areas'') are only recruited during difficult comprehension, or that right hemisphere neural activity does not contribute to the comprehension of discourse.

In addition, the lack of unique neural activation for high working memory capacity comprehenders when processing unpredictable inferences suggests that both high and low working memory capacity comprehenders may process unpredictable inferences in a similar manner. If a text provides little or no hints about an upcoming inference, comprehenders might not attempt to make a causal connection until they are forced to do so (i.e., when a coherence break occurs), regardless of their working memory capacity. For these unpredictable inferences, the coherence break cannot be resolved using semantic information previously primed by the text, so both high and low working memory capacity comprehenders likely have to search for causal relations before generating the necessary inference to successfully understand the text. This idea is consistent with findings showing longer reading times and poor memory for unrelated sentence pairs as compared to related sentence pairs (Myers et al., 1987). Thus, greater cognitive resources, such as a high working memory capacity, may not necessarily help comprehenders when they are not given any cues in the text and must later resolve an unexpected break in causal coherence.

These findings lead to several new issues about exactly how these neural substrates and cognitive processes interact during story comprehension. First, further work may shed light on the specific subprocesses involved when people generate these inferences. Such fractionation is beyond the scope of this study, although our theoretical framework (Jung-Beeman, 2005) suggests that the posterior temporal areas are involved in semantic activation and the inferior frontal areas are involved in semantic selection. Although the results of the current study did not find evidence to differentiate between these specific subprocesses, our findings do show important lateralization differences during text comprehension (e.g., bilateral activity for highly predictable text conditions and left hemisphere activity for less predictable text conditions).

Second, it is possible that moderately constrained text could differentially influence neural activity during inference generation. In the current study, the unpredictable text provides no information or hints about the upcoming coherence break. However, situations often arise in which a story provides some information or clues about an upcoming bridging inference (i.e., a moderately predictable text), without suggesting the event as strongly as our highly predictable text. It is possible that with moderately predictable text, differences in neural activity would be evident not only at the coherence break but also earlier in the inference generation process (perhaps when comprehenders generate predictive inferences).

Third, it is possible that these working memory capacity effects could differ throughout the time course of inference generation. Low working memory comprehenders could, in fact, show the same neural activity as high working memory comprehenders, only at later points during story comprehension. Although we did not show evidence of this pattern in our results, it would be interesting to investigate this issue with participants who have more extreme scores on a working memory capacity measure. In addition, it would be informative to examine working memory capacity effects more closely within each group of comprehenders. For example, some low working memory comprehenders may generate bridging inferences at different times than other low working memory comprehenders during story comprehension.

In conclusion, we observed distinct patterns of neural activity when people draw bridging inferences during story comprehension when these inferences are highly predictable in a text versus when these inferences are unpredictable in a text. Areas within both hemispheres (e.g., bilateral superior temporal gyri and inferior frontal gyri) contribute when comprehenders generate bridging inferences under highly predictable text conditions, whereas areas within the left hemisphere (i.e., the left middle and inferior gyrus) more strongly contribute when comprehenders generate bridging inferences under less predictable text conditions. In addition, these results show evidence that activity in the right hemisphere may reflect individual differences during story comprehension. Interestingly, these individual differences were only found under highly predictable text conditions. These findings suggest that a comprehender's working memory capacity may only influence text processing when some cues about an upcoming inference are provided in a text (i.e., under highly constrained text conditions). Thus, this study demonstrates a complex interplay between the level of predictability in a text (i.e., the causal constraint) and an individual's working memory capacity during story comprehension.

Acknowledgments

This research was supported through a grant from the National Institute on Deafness and Other Communication Disorders (DC-04052) to Mark Jung-Beeman.

We thank Zoe Clancy, Jason Haberman, and Therese Swan for their assistance in collecting data.

Reprint requests should be sent to Sandra Virtue, Department of Psychology, 2219 North Kenmore Avenue, DePaul University, Chicago, IL 60614, or via e-mail: svirtue@depaul.edu, or to Mark Jung-Beeman, Department of Psychology and Cognitive Brain Mapping Group, Northwestern University, 2029 Sheridan Road, Evanston, IL 60208, or via e-mail: mjungbee@northwestern.edu.

REFERENCES

Barch, D. M., Braver, T. S., Sabb, F. W., & Noll, D. C. (2000). Anterior cingulate and the monitoring of response conflict: Evidence from an fMRI study of overt verb generation. *Journal of Cognitive Neuroscience, 12,* 298–309.

Beeman, M., Bowden, E. M., & Gernsbacher, M. A. (2000). Right and left hemisphere cooperation for drawing predictive and coherence inferences during normal story comprehension. *Brain and Language, 71,* 310–336.

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Rao, S. M., & Cox, R. W. (1999). Conceptual processing during the conscious resting state: A functional MRI study. *Journal of Cognitive Neuroscience, 11,* 80–93.

Bryden, M. P. (1982). *Laterality: Functional asymmetry in the intact brain*. New York: Academic Press.

Collins, D. L., Neelin, P., Peters, T. M., & Evans, A. C. (1994). Automatic 3D intersubject registration of MR volumetric data

in standardized Talairach space. *Journal of Computer Assisted Tomography, 18,* 192–205.
Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, an International Journal, 29,* 162–173.
Estevez, A., & Calvo, M. G. (2000). Working memory capacity and time course of predictive inferences. *Memory, 8,* 51–61.
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. *Journal of Cognitive Neuroscience, 17,* 724–739.
Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101,* 371–395.
Jung-Beeman, M. (2005). Bilateral brain processes for comprehending natural language. *Trends in Cognitive Sciences, 9,* 512–518.
Jung-Beeman, M., Bowden, E. M., Haberman, J., Frymiare, J. L., Arambel-Liu, S., Greenblatt, R., et al. (2004). Neural activity when people solve verbal problems with insight. *Public Library of Science Biology, 2,* 500–510.
Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99,* 122–149.
Linderholm, T. (2002). Predictive inference generation as a function of working memory capacity and causal text constraints. *Discourse Processes, 34,* 259–280.
Linderholm, T., & van den Broek, P. (2002). The effects of reading purpose and working memory capacity on the processing of expository text. *Journal of Educational Psychology, 94,* 778–784.
Mason, R. A., & Just, M. A. (2004). How the brain processes causal inferences in text: A theoretical account of generation and integration component processes utilizing both cerebral hemispheres. *Psychological Science, 15,* 1–7.
McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review, 99,* 440–466.
Myers, J. L., Shinjo, M., & Duffy, S. A. (1987). Degree of causal relatedness and memory. *Journal of Memory and Language, 26,* 453–465.
Posner, M. I., Petersen, S. E., Fox, P. T., & Raichle, M. E. (1988). Localization of cognitive operations in the human brain. *Science, 240,* 1627–1631.
Seger, C. A., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (2000). Functional magnetic resonance imaging evidence for right-hemisphere involvement in processing unusual semantic relationships. *Neuropsychology, 14,* 361–369.
Singer, M., Andrusiak, P., Reisdorf, P., & Black, N. L. (1992). Individual differences in bridging inference processes. *Memory & Cognition, 20,* 539–548.
Singer, M., & Ritchot, K. F. M. (1996). The role of working memory capacity and knowledge access in text inference processing. *Memory & Cognition, 24,* 733–743.
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.
St. George, M., Mannes, S., & Hoffman, J. E. (1997). Individual differences in inference generation: An ERP analysis. *Journal of Cognitive Neuroscience, 9,* 776–787.
Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. New York: Thieme.
Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences, U.S.A., 94,* 14792–14797.
van den Broek, P. (1990). The causal inference maker: Towards a process model of inference generation in text comprehension. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 423–445). Hillsdale, NJ: Erlbaum.
Virtue, S., Haberman, J., Clancy, Z., Parrish, T., & Jung Beeman, M. (2006). Neural activity of inferences during story comprehension. *Brain Research, 1084,* 104–114.
Virtue, S. M., van den Broek, P., & Linderholm, T. (2006). Hemispheric processing of inferences: The effects of textual constraint and working-memory capacity. *Memory & Cognition, 34,* 1341–1354.
Wagner, A. D., Paré-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. *Neuron, 31,* 329–338.