NeuroImage 71 (2013) 59–74

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Cortical systems for local and global integration in discourse comprehension

Giovanna Egidi [a,*], Alfonso Caramazza [a,b]

[a] Center for Mind/Brain Sciences (CIMeC), University of Trento, Via delle Regole 101, 38123 Mattarello TN, Italy
[b] Cognitive Neuropsychology Laboratory, Harvard University, 33 Kirkland St., Cambridge, MA 02138, USA

ARTICLE INFO

Article history:
Accepted 5 January 2013
Available online 11 January 2013

Keywords:
Language processing
Brain–behavior correlation
Semantic memory
Context
Language network
Individual differences

ABSTRACT

To understand language, we integrate what we hear or read with prior context. This research investigates the neural systems underlying this integration process, in particular the integration of incoming linguistic information with local, proximal context and with global, distal context. The experiments used stories whose endings were locally consistent or locally inconsistent. In addition, the stories' global context was either relevant or irrelevant for the integration of the endings. In Experiment 1, reading latencies showed that the perceived consistency of an ending depended on its fit with the local context, but the availability of a relevant global context attenuated this effect. Experiment 2 used BOLD fMRI to study whether different neural systems are sensitive to the local consistency of the endings and the relevance of the global context. A first analysis evaluated BOLD responses during the comprehension of story endings. It identified three networks: one sensitive to consistency with local context, one sensitive to the relevance of the global context, and one sensitive to both factors. These findings suggest that some regions respond to the holistic relation of local and global contexts while others track only the global or the local contexts. A second analysis examined correlations between BOLD activity during listening of the story endings and subsequent memory for those endings. It revealed two distinct networks: Positive correlations in areas usually involved in semantic processing and memory for language, and negative correlations in sensory, motor, and visual areas, indicating that weaker activity in the latter regions is conducive to better memory for linguistic content. More widespread memory correlates were found when global context was relevant for understanding a story ending. We conclude that integration at the discourse level involves the cooperation of different networks each sensitive to separate aspects of the task, and that integration is more successfully achieved when the processing of potentially distracting information is reduced.

Introduction

Discourse comprehension is a process of progressive meaning construction that entails dynamic integration of incoming information with local, proximal context and with global, distal context (Kintsch, 1988, 1998). The local context is the most recently encountered information, and the global context is information introduced earlier. While reading or listening to a story, for example, people treat as local context for an incoming sentence information introduced in the prior one or two sentences, and treat as global context information introduced in earlier portions of the text (e.g., Albrecht and O'Brien, 1993; McKoon and Ratcliff, 1992).

Psychological theories of discourse comprehension, drawing on behavioral evidence and computational models, have often highlighted the importance of understanding how the balance between integrating local and global information is reached at the cognitive level. Theories differ in their view of the mechanisms and situations that bring global contexts to the integration process, but all agree that incoming information is always integrated with local information and with some portions of the global context (for a review, see McNamara and Magliano, 2009). Thus, local integration is considered to be the default process in discourse comprehension, because local information is always available in working memory. The possibility for global integration, however, is determined by potential access to distal, less-recently encountered information in working memory at the time integration occurs. This access depends on the extent to which features of the incoming information function as memory cues for distal information, and on the extent to which certain types of information (e.g., time, space) are continually monitored and kept active (Gerrig and O'Brien, 2005; McNamara and Magliano, 2009; Zwaan and Radvansky, 1998). Our goal in this paper is to understand how distal, global context modifies the integration of incoming information with proximal, local context.

The current study presents the first examination of the neurobiological networks mediating local and global integration during comprehension. To date, neuroimaging studies of discourse comprehension have not examined these two types of integration separately. They have either only examined local integration, or they have not distinguished between contexts requiring local and global integration. These studies

* Corresponding author. Fax: +39 0461 283066.
E-mail address: giovanna.egidi@unitn.it (G. Egidi).

1053-8119/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2013.01.003

have identified general activity patterns associated with comprehending passages as compared to rest (e.g., Kansaku et al., 2000; Papathanassiou et al., 2000), to reversed speech (e.g., Crinion and Price, 2005; Crinion et al., 2003; Kansaku et al., 2000), and to unconnected sentences (e.g., Xu et al., 2005; Yarkoni et al., 2008). Other studies have examined coherence breaks in discourse, where the distance between the critical sentence and the text that created the coherence break was either small within the local context (Hasson et al., 2007) or was not controlled (Deen and McCarthy, 2010; Ferstl et al., 2005). While all these methodologies certainly highlight central features of discourse integration, they were not designed to investigate the neural activity associated with local and global integration separately.

Those neuroimaging studies that have focused on distinguishing between local and global integration have used the term global integration in a different sense than we do here—not in terms of integration with prior linguistic content provided in the text, but as integration with general prior knowledge stored in long-term memory (Maguire et al., 1999; Martín-Loeches et al., 2008; Menenti et al., 2009; St George et al., 1999). Most of these studies followed a classic paradigm in psycholinguistics (Bransford and Johnson, 1972): Participants read an ambiguous passage which was preceded or not by an explanatory context in the form of either a title or a picture (Maguire et al., 1999; Martín-Loeches et al., 2008; St George et al., 1999). The passages used in those studies were difficult to understand, to the point that even local consistency was disrupted when the explanatory context was not given (e.g., *Once you are settled, your thumbs should be pointing up. Sometimes there is no security but the animal's hair*; Martín-Loeches et al., 2008; St George et al., 1999). These studies too, in most cases, did not examine local and global integration (as we intend it) separately. Both global and local coherence were either totally disrupted when no framing title or picture was given, or both could be achieved when the general topic was disclosed in advance. Thus, the processes of global and local integration were overlapping and the brain regions associated with the two processes could not be assessed separately.

Thus, prior work has left open the fundamental question of whether integration of incoming sentences with local and global contexts involves a network mediating integration with both kinds of contexts, or whether the integration with local and global contexts relies on partially different systems. To address this issue, we devised a paradigm that kept local and global contexts distinct so that it would be possible to differentiate their relation to incoming information and identify the brain regions associated with local and global integration separately. To this end, we created narratives that could have four versions (see Fig. 1 for an example). The first few sentences of each story formed the global, distal context, and the last few sentences the local, proximal context for a final sentence, which was the critical test sentence. This ending sentence was manipulated to always be either consistent or inconsistent with the local context. Independently, the global context was manipulated in a way that best allowed testing the impact of distal information on the integration process: It was either irrelevant or relevant for the integration of the ending. When the global context was irrelevant, its presence in the story had no impact on the integration of the endings, and local consistency was the only determining factor, because the integration of the ending depended solely on the local context. However, when the global context was relevant, there was always a possibility for consistency: The ending that was inconsistent with the local context was in this case consistent with the global context. Conversely, the ending that was consistent with the local context became inconsistent with the global context. The rationale for creating a situation of conflict between the demands of the local and global contexts was to dissociate their impact on the integration process. This opposition was crucial, as it allowed us to test how the global context can prevail on the requirements of the local context when it is informative during the integration of the endings.

The logic of our design shares commonalities with ERP studies on discourse comprehension that measure detection of inconsistencies as modulations of the N400 effect. These studies examine contextual factors that have direct bearing on meaning construction of a sentence and test whether global context can override integration constraints at the sentence level. For example, the word *slow* in the sentence *Jane told her brother that he was exceptionally slow* evokes a greater N400 than *quick* in *Jane told her brother that he was exceptionally quick* if the sentence is presented after a context in which the brother is described as being quick. This occurs even though at the sentence level both adjectives are equally coherent (van Berkum et al., 1999, 2003). Analogously, other studies have shown that different features of global discourse context can outweigh the detection of semantic violations within a sentence and elicit smaller N400 for words that match the global but not the sentence context (e.g., Nieuwland and Van Berkum, 2006). More generally, these ERP studies show that several aspects of the global context can modulate word-by-word incremental integration as indexed by the amplitude of the N400 (for a review, see Hagoort and van Berkum, 2007). Our own research has also shown that the N400 peak latency is delayed when understanding the implication of a story ending requires global (as opposed to sentence) integration (Egidi and Nusbaum, 2012).

The design of the current research, which strongly opposed local and global integration, allowed us to identify three functional networks:

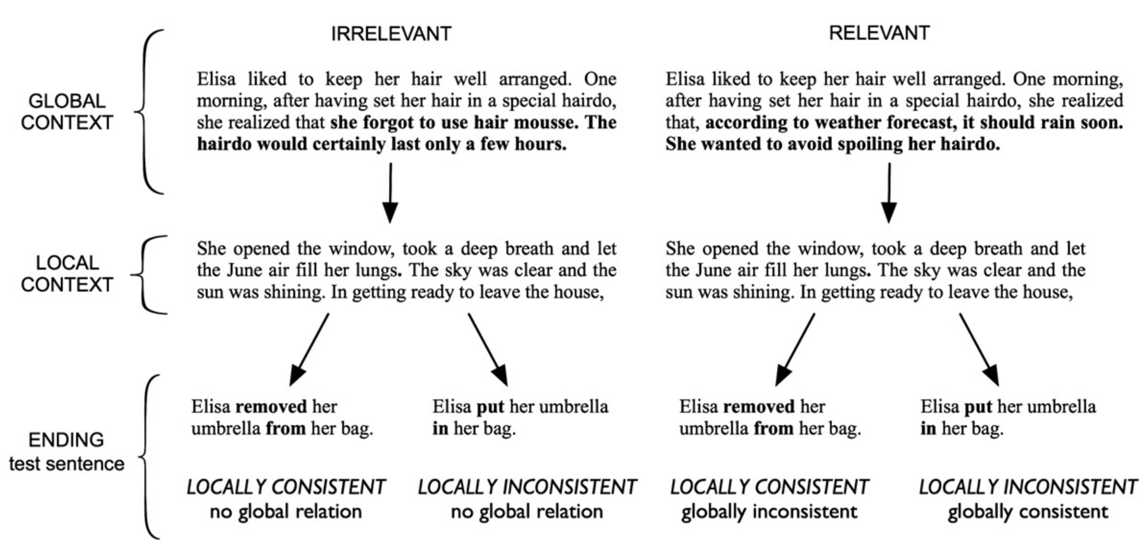

Fig. 1. Example of stimulus and design structure. Sample story used in the experiments and schematic of the experimental conditions.

Two associated with different aspects of global integration and one with local integration. The first network for global integration includes the regions in which the BOLD response during the comprehension of the story endings differs as a function of the relation between these sentences and both the local and global contexts (the interaction). This network is associated with readiness to the information that makes most sense given the full context. Several theories of discourse comprehension argue that comprehenders (in particular readers) routinely attempt to construct a meaningful representation of a text that is coherent overall (e.g., Graesser et al., 1994; Long and Lea, 2005). For this to occur, a certain amount of top–down processing is necessary, resulting in easier integration of information generally consistent with prior discourse. In our experiment, we constructed our stimuli so that sensitivity to both local and distal contexts would be reflected in greater receptiveness to the most sensible endings in the context of the entire story. We therefore hypothesized that the regions associated with this global integration would be those identified in prior work as sensitive to top–down attentional processes and to processes of goal-directed selection for stimuli. These include dorso-parietal cortex along the intraparietal sulcus extending to the superior parietal lobule, as demonstrated in studies on memory retrieval and visual attention (e.g., Cabeza et al., 2008; Corbetta and Shulman, 2002).[1]

The second network, also crucial for global integration, includes the regions in which the BOLD response during the comprehension of the endings differs as a function of the relevance of the global contexts. This system is associated with accessing distal information. We reasoned that because access to distal information entails holding a greater amount of information, it would entail a greater working memory load. Although our method cannot tell what is the content of working memory at the time participants listened to the endings, we based this prediction on the premise that the possibility for global integration is determined by access to distal information during integration, as shown by several psychological studies of narrative comprehension (Gerrig and O'Brien, 2005; McNamara and Magliano, 2009; Zwaan and Radvansky, 1998). We therefore expected that regions associated with working memory maintenance would show greater BOLD response for the integration of endings with relevant global contexts than for the integration of endings with irrelevant contexts. These regions include aspects of the dorsolateral and ventrolateral prefrontal cortices and of the parietal lobules (for reviews, see Rawley and Constantinidis, 2009; Wager and Smith, 2003).

The third network our design allows us to identify is that of local integration. We reasoned that inconsistencies at the local level should be very easy to detect, and therefore we expected a pattern reflecting detection of inconsistency. An inconsistency detection system has been identified in several studies on language comprehension (not only discourse) involving inferior frontal gyrus, dorsomedial and ventromedial prefrontal cortex, superior temporal sulcus, the middle part of the middle temporal gyrus, anterior temporal lobe, and posterior midline regions (Ferstl et al., 2008; Hasson et al., 2007). In this network, inconsistent information is usually associated with increased BOLD response with respect to consistent information. In our experiment, therefore, this network should show increased BOLD response for locally inconsistent endings.

Our experiment was designed to examine also a second issue. After listening to all the stories in the fMRI scanner, participants carried out a surprise recall test that assessed their memory for the stories (outside the scanner). We added this task to identify the brain regions associated with the cognitive processes mediating successful comprehension of the endings, when these reflected either local or global integration. This was achieved by correlating neural activity associated with listening to the story endings with subsequent recall scores for these endings. Similar analyses have been performed in prior studies (Crinion and Price, 2005; Hasson et al., 2007; Maguire et al., 1999; Yarkoni et al., 2008) and have proven valuable for understanding the relationship between brain activity during integration and the behavioral output that is the result of that integration process. Prior studies have identified correlations with frontal areas such as prefrontal cortex, inferior and superior frontal gyri (Crinion and Price, 2005; Hasson et al., 2007; Maguire et al., 1999; Yarkoni et al., 2008), parietal areas such as inferior parietal lobule (Hasson et al., 2007), temporal areas such as fusiform gyrus, posterior-dorsal superior and medial temporal gyri and midline regions such as posterior cingulate gyrus and medial prefrontal cortex (Hasson et al., 2007; Yarkoni et al., 2008). Our analysis differs from those performed in prior studies in that it correlates memory scores with the cortical activity of each experimental condition. In our experiment, we expected that when global context is relevant for the integration of the ending, memory for the ending would depend on a more extended brain network than when global context is irrelevant, because the comprehension of the ending following a relevant global context requires handling a greater amount of information and makes the integration task more demanding.

We began our examination of local and global integration processes with a behavioral experiment based on self-paced reading. This study tested the impact of global and local contexts on the comprehension of story endings and was a necessary precursor to the neuroimaging study because it established that, with our materials, people were sensitive to both the local and global contexts during story comprehension. Based on prior literature (e.g., Egidi and Gerrig, 2006; Magliano and Radvansky, 2001), we expected that reading latencies would show readers' strong tendency for local integration, as seen by longer latencies for locally inconsistent endings. However, we also expected that this effect would be modulated by sensitivity to the global context and thus be reflected in a slowdown for locally consistent endings when the story contained relevant global context (that made those endings globally inconsistent), and in a speed-up for locally inconsistent endings when the story contained relevant global information (that made those endings globally consistent). In short, we expected that the effect of local consistency would be moderated by global consistency.

Materials and methods

Participants

Thirty naive volunteers participated in the behavioral study (Experiment 1) and fourteen other participants in the fMRI study (Experiment 2). All participants were native Italian speakers. Participants in the fMRI experiment (M age = 25.29; SD = 4.86; f = 4) were right-handed, had good or corrected vision, good hearing, did not report mood or attention disorders, and did not take psychotropic medications. Prior to the scan they underwent a medical interview, which evaluated other exclusion criteria. The ethical committee for research involving human subjects at the University of Trento approved the experimental procedures.

Stimuli

The stimuli consisted of 40 stories in Italian, 6 to 10 sentences long, describing simple, everyday life events. We were interested in integration processes occurring while people comprehended the last sentence of the stories, which described a protagonist's action. With respect to this story ending, the first few sentences of each story constituted a wide context (global context), and the following few sentences constituted a narrow context (local context). Fig. 1 shows a sample story and a schematic depiction of the experimental conditions. Additional stories are available in Supplementary Appendix A in the Supplementary materials.

[1] The network also includes frontal eye fields, but their activation is due to visual processing in visual tasks (cf. Corbetta and Shulman, 2002). Because in our experiment presentation is auditory, we do not expect these regions to be involved and we will not discuss them further.

Each story could have two endings; one consistent and one inconsistent with the local context. There were four experimental conditions. In two conditions, the global context was irrelevant to the integration of the ending; in those cases, the coherence of the ending with prior content depended only on the local context. In the other two conditions the global context was relevant for the integration of the ending. In those cases, a locally consistent ending was *in*consistent with the global context and, conversely, a locally *in*consistent ending was consistent with global context. A norming study validated that the locally consistent and locally inconsistent endings were indeed judged as such by participants and that global context moderated these effects. Norming procedure and results are reported in Supplementary Appendix B, Supplementary Figure A, and Supplementary Table A.

For each story, we balanced lexical and semantic overlap across versions. All ending sentences were between 12 and 14 syllables long, and had similar syntactic structure. We took these precautions so that it would be possible to compare both reading latencies and BOLD activity associated with the comprehension of the endings in the behavioral and fMRI experiments. We used a Latin square to assign the stories to different lists in a counterbalanced fashion, so that each participant would be presented with only one version of each story (with one type of global context and one type of ending). This resulted in the construction of 4 lists in each experiment, each containing 40 experimental stories. Lists were different for each experiment, and stories were assigned randomly to each list.

In addition to the main experimental materials, we wrote 20 stories similar in length and structure to the experimental narratives, which we used as practice materials in the two experiments and as fillers in the fMRI experiment. Stories were presented visually in the behavioral experiment and auditorily in the fMRI experiment. The speaker who recorded the stories for the fMRI study was blind to the purpose of the experiment, and recorded the story endings separately from the stories, so as not to bias intonation in reading the story bodies as a function of the endings. In the fMRI study we also used two video clips of about 8 min each, so that participants would not experience discomfort in the scanner or fall asleep during periods in which they were not required to perform any experimental task.

Design

The design was the same in both behavioral and fMRI experiments. The stories presented could have an irrelevant or a relevant global context for the integration of the endings, and the endings could be either consistent or inconsistent with respect to the local context. Thus, the design consisted of two within-participants variables: Global Context Relevance (Irrelevant, Relevant) and Ending Local Consistency (Consistent, Inconsistent). Throughout this report we refer to the four experimental conditions as Irr_Cons, Irr_Inco, Rel_Cons and Rel_Inco, where Irr and Rel refer to the relevance of the global context and Cons and Inco to the local consistency of the ending sentences. This design allowed us to evaluate whether participants integrate the endings with global context when this is relevant and whether global context overrides or interacts with the effects of the more recently presented information appearing as local context.

Behavioral procedure in Experiment 1

Participants read each story one line at a time in a self-paced manner. They were told to move to the next line only once they had comprehended the line currently on the screen. Each story was followed by a comprehension question with three possible answers: *Yes*, *No*, *Cannot tell from the story*. The questions were designed so that each of these answers was correct for about 33% of the stories. At no point during the experiment were participants allowed to see earlier screens. Stories were presented in a different random order for each participant. Participants were instructed to read each line for comprehension, as naturally as possible. They were also asked to answer the comprehension questions as accurately as possible, grounding their answer on the information contained in the story they had just finished reading. Before starting the experiment, participants practiced the procedure on three filler stories.

Behavioral data processing and statistical analyses in Experiment 1

We discarded data from two participants whose accuracy in responding to the comprehension questions was below 75%. We then removed responses that took longer than 2.5 standard deviations above and below each condition's group mean. This procedure resulted in a loss of 3% of the data. We conducted an omnibus 2 (Global Context Relevance) $\times$ 2 (Local Ending Consistency) ANOVA on mean reading latencies and four planned contrasts that compared the two Local Consistency conditions for each Context Relevance level and the two Context Relevance conditions for each Local Consistency level.

fMRI and behavioral procedure in Experiment 2

The experiment consisted of two phases: A phase in which we acquired fMRI data while participants listened to the experimental materials, and a behavioral testing phase outside the scanner in which we asked participants questions on the stories and their experience in the scanner. The experiment included the 40 experimental stories and 12 filler stories, which were very similar in structure and type of contents. Fig. 2 shows a schematic of the procedure for both experimental and filler stories. To ensure that participants would process both types of stories in a similar manner, we used the same procedure for their presentation. Differences between stories occurred only once the stories had been presented in their entirety: Filler stories were followed by a comprehension question requiring a yes/no response by a button press, and experimental stories were followed by a visual stimulus indicating that a button press was required. We used comprehension questions to ensure that participants would pay attention to the content of the stories throughout the experiment, and button presses to ensure that participants would follow the progress of the experiment. We chose not to ask comprehension questions on experimental trials because we did not want to contaminate the hemodynamic response of the experimental items with the processing associated with answering to a question. In addition we did not want to compromise the recall test that followed scanning (see details below) by giving participants the chance to rehearse the content of the stories. When asked during debriefing, participants reported that while listening to the stories they were not able to predict whether a trial would end with a button press or a question. In addition, since filler stories were slightly shorter than experimental stories, we established in a pilot study that this difference was not noticeable. The task was listening for comprehension, as we wanted to study comprehension in the most natural way, thus minimizing the chance that participants would engage in any strategic processing.

Before participants entered the scanner, they were given instructions and practiced the procedure on a block of 4 filler stories. Trials were described as follows (in Italian, translation follows): "Each trial begins with a warning sound and the appearance of a cross in the center of the screen, followed by a voice telling a story. Please pay attention to the story as to a friend's voice recounting an anecdote. At the end of each story the cross disappears from the screen and you will be asked to do one of two things. In some cases, a question about the content of the story appears on the screen. You can answer either *Yes* or *No* to the question, by pressing one of the two keys on the button box. You must do so within 4 s, because after this delay the question disappears from the screen and the software will not be able to record your response. In other cases, after hearing the story you will see two circles on the screen. These too are displayed for

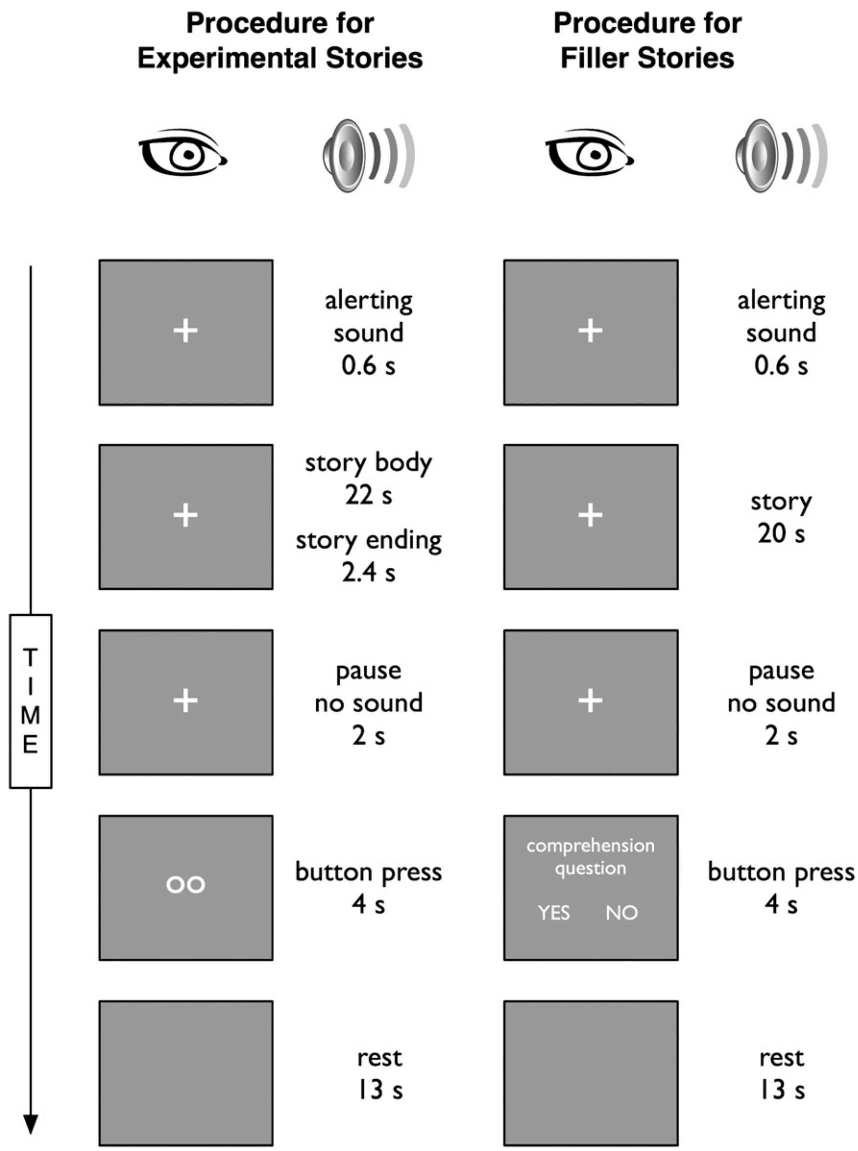

Fig. 2. Procedure for auditory story presentation in the MR scanner (Experiment 2). The procedure was identical during the presentation of experimental and filler stories. It differed only after: Comprehension questions followed the presentation of filler stories to verify that participants listened to the content of the stories; button presses followed the presentation of the experimental stories to maintain vigilance over the study. Critical window: 4.4 s between the onset of the ending sentence and the onset of the button press screen.

only 4 s and you should press either key on the button box in this time period. When this or the question screen disappears, you will see a gray screen for several seconds. During this time your only task is to rest." Participants responded using a two-key button box. Half of the participants used their left hand, and half the right hand.

We also informed participants of the structure of the experiment: We told them that they would watch a video clip for about 8 min, followed by two blocks of stories of 10 min each, followed by another video clip for 8 min, and two more blocks of stories of 10 min each. The video clips were presented during the acquisition of the structural images. Participants were asked to pay attention to the video clips and stories equally, and were told that they would be asked questions about both once they finished the part in the scanner.

The second phase of the study took place outside the scanner and consisted of a recall test assessing participants' memory of the video clip and story endings. They first described the general subject and the location where the events of each video clip took place. Next, in a cued recall paradigm, participants were presented with the text of each of the stories and were asked to write how each story ended. We used these memory protocols to create recall scores against which neural activity during the study was correlated. As a final task, participants completed a survey about the experiment they had just completed and some personal preferences. Because the recall data of the video clips and the data of the final surveys are not relevant for the questions addressed here, we will not discuss them further.

fMRI data acquisition in Experiment 2

We acquired fMRI data in a Bruker MedSpec 4 T scanner with an eight-channel head coil at the Center for Mind/Brain Sciences (CIMeC) of the University of Trento. We collected functional scans with EPI sequences, which consisted of 37 axial T2-weighted functional images. They were acquired every 2200 ms in ascending interleaved order, and were slightly tilted to run parallel to the AC–PC line. We used isotropic voxels of 3 mm, echo time of 33 ms, flip angle of 75°, field of view of 192×192 mm, and gap size between slices of .45 mm. Participants completed 4 runs of 270 volumes each. Before each functional run, we performed an additional brief scan that measured the point-spread function (PSF) of the acquired sequence and whose purpose was to correct for possible distortions due to the non homogeneity of the magnetic field in certain regions (Zaitsev et al., 2004).

To superimpose functional activation maps on anatomical scans, we also acquired high-resolution anatomical images with a T1-weighted MP-RAGE sequence. These images consisted of 176 sagittal images, acquired every 2700 ms, with isotropic voxels of 1 mm, echo time of 4 ms, flip angle of 7°, and field of view of 256 × 254 mm. We acquired two anatomical images, one at the beginning of the experiment, and one after participants had completed two experimental runs.

fMRI data processing, surface projection, and statistical analyses in Experiment 2

We performed all functional data analyses by using AFNI's procedures (Cox, 1996). For each participant's functional data, we first removed the initial 11 volumes of each run, which we had acquired before the beginning of the experimental task to allow for scanner stabilization. After registration of all runs to a reference time point in the first run, we performed motion correction, applied a Gaussian blur with a full width at half maximum (FWHM) of 6.0 mm, and mean-normalized the time series to obtain percent signal change values.

Analysis of time series was done using simultaneous regression implemented using AFNI's 3dDeconvolve regression routines. Regressors were waveforms with similarity to the hemodynamic response. These regressors were generated by convolving a gamma-variant function with the onset time and duration of the endings and with the onset time and duration of the stories. There were five regressors of interest, one for each experimental condition (the four types of endings) and one that modeled the portion of the story that preceded the ending together with the filler stories. The regression solution returned a Beta value for each regressor. These Beta values were based on partial correlations modeling whether the presentation of each of the elements of interest accounted for variance above and beyond the variance associated with the other events occurring in the experiment.

We modeled all story bodies with a single regressor, on the assumption that story content was quite similar up to the final sentences; that is, although of different relevance for the ending sentences, global contexts would not have an impact on how participants processed the local context. However, it can be argued that, because irrelevant and relevant global contexts consist of different texts, they should be modeled separately. We therefore performed an additional analysis (of the BOLD response during comprehension only) that used separate regressors for irrelevant, relevant, and filler story bodies. Since the results of these two analyses are very similar, we report and discuss in the manuscript the results of the analysis that used only one regressor for all story bodies. The results of the other analysis are available in Supplementary Figure B and Supplementary Table B.

Other regressors reflected factors of no interest and included the button presses, the comprehension questions to the filler stories, the 6 motion parameters estimated during head motion correction, and 1st–4th order polynomial trends fitted for each run separately to account for instrumentation-induced drifts in the signal. We also removed from the regression functional acquisitions associated with strong head movement (> 2 mm) and acquisitions that presented outlier values in a large number of voxels (> 2000). This removal accounted for 4.8% of the data.

For each participant, we co-registered and averaged the two MP-RAGE images, to increase the image SNR. We then inflated left and right hemispheres of each participant to a surface representation and we aligned them to a common template using warping procedures implemented in FreeSurfer (Fischl et al., 1999). Then, by averaging participants' individual cortical surfaces, we created a mean cortical surface to be used for the group analysis of functional data. Finally, we projected all results to this average cortical representation for display purposes. For each participant we aligned the participant's anatomical image to the EPI data with procedures implemented in AFNI (Saad et al., 2009). We projected the results of the regression analysis from the 3D volumes to the 2D cortical surfaces.

We conducted two second-level group analyses. The first identified brain regions where activity for story endings varied as a function of the two experimental variables and their interaction. We conducted a whole brain vertex-wise analysis in the 2D surface domain where for each vertex, we analyzed the Beta values for all participants (modeled as a random factor) using a 2 (Global Context Relevance) × 2 (Ending Local Consistency) ANOVA. The second identified cortical regions where inter-individual differences in BOLD response correlated with accuracy on the recall test. Because participants listened to four types of content, we conducted four correlation tests for each vertex in the 2D surface domain; that is, one correlation between responses to each of the four experimental conditions (Irr_Cons, Irr_Inco, Rel_Cons, Rel_Inco) and the memory for each ending given the relevance of the global context. This analysis produced four statistical maps identifying brain regions involved in memory for particular types of content.

All analyses were controlled for family-wise error (FWE) using cluster-based constraints: Cluster thresholding was based on Monte Carlo simulation methods (Forman et al., 1995) that controlled for Family-Wise Error (FWE) rate at $p < .05$. These simulations take into account the smoothing in the data, the allowed distance between active vertices (2 mm), and voxel thresholding. We used 2 single-vertex thresholds set at an alpha level of $p < .05$ and $p < .005$ to identify both larger, distributed clusters and smaller, focal clusters (as in e.g., Hasson et al., 2007). Note that using cluster-based thresholding for FWE correction means that cluster size is not chosen arbitrarily, but is tightly dependent on the voxel threshold selected (more liberal single-voxel thresholds result in large cluster-size magnitudes). The results of both clusterings are shown in the figures and tables. However, discussion of the results in the paper is mainly based on the clustering at a voxel threshold of $p < .05$, as the more focal clusters tended to be subsets of the larger ones. From the clustering at a voxel threshold of $p < .005$, only the clusters that do not overlap with those of clustering at a voxel threshold of $p < .05$ are explicitly discussed.

Behavioral data processing and statistical analyses in Experiment 2

After the fMRI part of the study, participants completed a recall task of the story endings they had heard in the scanner. We divided each ending into meaningful units capturing an idea that could be easily remembered as a whole. Two independent raters, blind to the global context relevance conditions, assigned one point to each idea fully recovered, half a point to ideas partially recovered and zero points to missing ideas, misrecalls, and guesses. For each meaningful unit, raters also assigned one point for perfect verbatim memory, and half for partial. The two gist and verbatim scores were then added together and transformed into a percentage score for each rater. On the combined recall scores thus obtained, we conducted an omnibus 2 (Global Context Relevance) × 2 (Local Ending Consistency) ANOVA and a comparison of the two endings when the global context was irrelevant. We did this for the scores of each of the two raters separately. Because the results of these statistics were the same and the inter-rater reliability was very high (Cronbach $\alpha = .99$), we used the scores of one rater (those reported in Fig. 4) for the BOLD-to-memory correlations.

Behavioral results

Experiment 1: self-paced reading of story endings

The purpose of Experiment 1 was to evaluate the potential impact of global context on online integration processes. This was done by measuring reading latencies for story endings in the four experimental conditions. As prior research has shown (e.g., Egidi and Gerrig, 2006; Magliano and Radvansky, 2001; McKoon and Ratcliff, 1992), there is

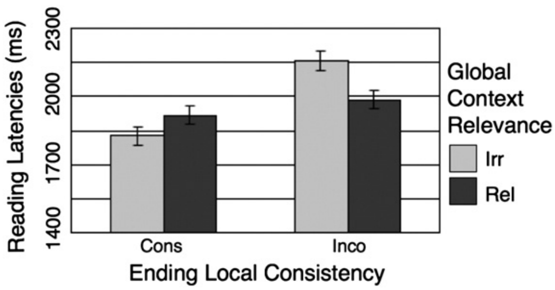

Fig. 3. Results of Experiment 1. Mean reading latencies for story endings as a function of local consistency and global context relevance. Confidence intervals for within-participant effects are calculated following Loftus and Masson (1994).

often a tendency for local integration during online processing. This experiment allowed us to test whether this result could also be obtained with our stimuli. Fig. 3 shows the pattern of results and Table 1 reports statistically significant effects. Participants read locally consistent endings more quickly than locally inconsistent endings, as shown by the main effect of local consistency in the omnibus ANOVA. However, this effect was modulated by an interaction with global context relevance; participants read locally consistent endings slower when the story contained relevant global information that made those endings globally *in*consistent. Similarly, participants read the locally *in*consistent endings faster when the story contained relevant global information that made those endings globally consistent. This interpretation is statistically supported by the planned contrasts reported in Table 1. Thus, participants demonstrated sensitivity to the global context, but also—and most importantly—a strong tendency to integrate information locally.

Experiment 2: recall of story endings

After the fMRI scanning session, participants completed a cued-recall test that evaluated their memory for story endings in each of the four experimental conditions. Behavioral evidence has shown that, when comprehenders have a high knowledge base on which to map the text (which, given the everyday-life type of events described in our stories can be safely assumed to occur in this experiment), memory for text is usually better for portions that are perceived as inconsistent (Kintsch, 1994; McNamara and Kintsch, 1996; McNamara et al., 1996). The explanation given for this result is that inconsistent information is given greater attention and more elaborate processing because its integration with prior context is more difficult (or altogether impossible). Our stimuli were constructed so that when the global context was irrelevant, the consistent ending would follow naturally from the story. Therefore, because this content is very easily integrated with prior context, it would be remembered less well than an inconsistent ending. When the global context was relevant for the integration of the endings, however, either ending could be perceived as inconsistent with a portion of the story, but could also be successfully integrated with a different part of prior context (either the local or the global context). This potential for integration permits the inclusion of the endings in a larger semantic structure, which in turn can lead to better memory. This reasoning yielded two predictions. The first prediction is that recall scores would reflect overall better memory when the global context was relevant, for both locally consistent and locally inconsistent endings. The second prediction is that memory would be better for locally inconsistent endings (as compared to locally consistent) when the global context was irrelevant.

Both predictions were confirmed in the study. Behavioral results for the recall task are shown in Fig. 4. Participants recalled more information about the story endings when the global context was relevant to their integration than when it was not. This result was corroborated by the main effect of relevance found by the omnibus ANOVA ($F(1,13) = 19.05$, $MSe = 0.42$, $p < .001$). Although this ANOVA did not reveal a reliable interaction, the contrast based on our specific a-priori predictions revealed that locally consistent and locally inconsistent endings were associated with different recall levels only when the global context was irrelevant ($F(1,13) = 8.19$, $MSe = 0.19$, $p < .05$). That is, when the story offered only one possibility for consistency, participants remembered locally inconsistent endings better than locally consistent ones. These results are consistent with Experiment 1's reading-time results and show that the experimental variables achieved the intended effects also during passive listening. These results also provide support for the recall model outlined above, as well as further validating the operationalization of local consistency and global context relevance in the current study.

fMRI results

On the fMRI data collected in Experiment 2, we first performed a preliminary analysis aimed at validating our basic results against prior literature on language comprehension. This analysis identified regions that showed an increase or a decrease in BOLD response relative to baseline during listening to the story bodies. The implicit baseline comprised the residual activity after all other events occurring in the experiment had been modeled (i.e., activity associated with listening to story endings, pressing the button, resting, and processing the comprehension questions). The results demonstrate a pattern consistent with prior neuroimaging research on spoken discourse comprehension with strong activation in bilateral temporal regions and left inferior frontal gyrus (IFG), as well as strong deactivation in regions associated with the default-mode network (Raichle et al., 2001). These results are shown in Supplementary Figure C and reported in detail in Supplementary Appendix C.

Two analyses addressed the theoretical questions that motivated the study. One identified the regions associated with different BOLD responses during processing of story endings as a function of the experimental variables—that is, (a) endings' consistency with their local context and (b) the relevance of the global context preceding the

Table 1
Effects of independent variables on self-paced reading latencies (Experiment 1). Reading latencies showed sensitivity to consistency of endings with local context, as well as an interaction with the relevance of global context.

Omnibus ANOVA	*F*	*MSe*	*p*<
Interaction	14.21	30,784.29	.001
Contrasts			
Irr Cons<Rel Cons	4.13	38,468.97	=.052
Irr Inco>Rel Inco	5.63	51,192.15	.05
Irr Cons<Irr Inco	29.93	52,850.02	.001
Rel Cons<Rel Inco	5.44	19,098.88	.05

dfs = 1,27.

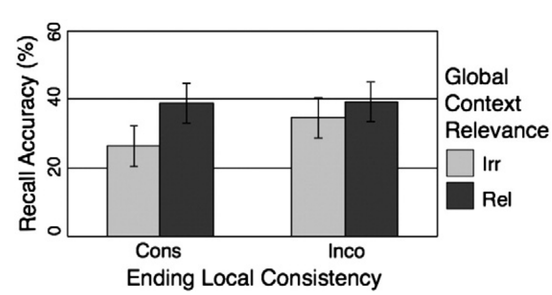

Fig. 4. Results of recall scores in Experiment 2. Percent recall accuracy and confidence intervals (Loftus and Masson, 1994) for the story endings as a function of the experimental variables. Recall scores combine gist and verbatim memory.

presentation of these endings. The other analysis identified regions in which inter-individual differences in neural activity predicted later memory by means of correlations between recall accuracy scores and participants' BOLD responses in these conditions.

Neural activity associated with comprehension as a function of the experimental variables

This analysis identified brain regions where BOLD responses during comprehension of story endings differed as a function of the experimental variables (local ending consistency and global context relevance). Behavioral evidence in the literature and the results of Experiment 1 show that comprehenders have a tendency to integrate narrative information locally (e.g., Egidi and Gerrig, 2006; Magliano and Radvansky, 2001; McKoon and Ratcliff, 1992) but they are also sensitive to the broader demands of the global context (e.g., Guéraud et al., 2005; Zwaan et al., 1995). This analysis was therefore aimed at identifying brain regions associated with local and global integration processes.

We probed for three groups of regions; those demonstrating the interaction of the two experimental variables and those demonstrating either of the two main effects. Specifically, the first set of regions included clusters of cortical areas sensitive to the interaction between the local consistency of the endings and the relevance of the global context. The second set included clusters where activity differed depending on whether the endings were preceded by irrelevant or relevant global contexts (global context relevance effect). The third set of regions included clusters of cortical areas where activity differed during the listening of locally consistent and locally inconsistent endings (local consistency effect). Fig. 5 and Table 2 show the results of these analyses.

Interaction

This analysis identified regions sensitive to the relation between the ending sentence and both the local and global contexts. One cluster was found in superior parietal regions in the left hemisphere, and included the intraparietal sulcus (IPS) and the lateral extension of the parieto-occipital sulcus (POS). One additional cluster encompassed the paracentral lobule and the dorsal central sulcus (CS) bilaterally, and on the left it also extended to the dorsal precentral gyrus (PreCG).

The pattern of activity for the parietal cluster showed above baseline activation for all conditions. When the global context was irrelevant, the activity for locally consistent endings was greater than the activity for inconsistent endings. In addition, the activity for locally

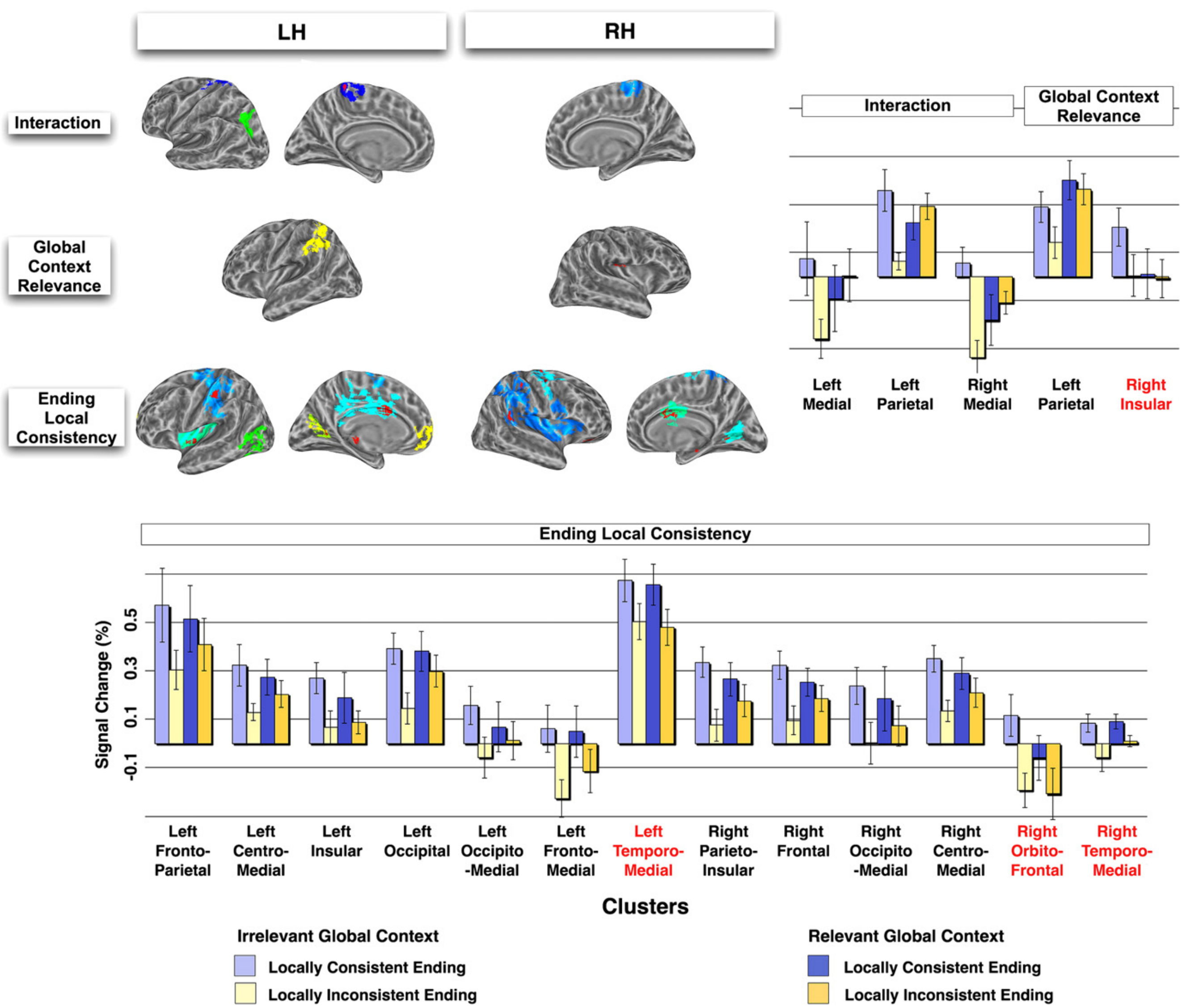

Fig. 5. Clusters of activity associated with the experimental variables in Experiment 2. BOLD activity associated with listening to the story endings in the four experimental conditions. Red clusters: p<.005 voxel threshold. LH = left hemisphere, RH = right hemisphere.

Table 2
Talairach coordinates of centers of clusters where BOLD activity varied as a function of ending local consistency, global context relevance, or their interaction (Experiment 2). Clusters match regions in Fig. 5: Numbers refer to large clusters (p<.05 voxel threshold) and letters to focal clusters (p<.005 voxel threshold).

Effect	Region (focus point)	Area (mm^2)	x	y	z	Cluster
Interaction	L precentral gyrus	840	−6	−30	62	1
	L intraparietal sulcus	765	−19	−60	33	2
	R paracentral lobule	569	17	−33	64	3
Global Context Relevance	L supramarginal gyrus	1196	−29	−37	36	1
	R insula	44	39	−19	8	A
Ending Local Consistency	L central sulcus	2933	−34	−29	60	1
	L cingulate gyrus	1764	−4	−22	28	2
	L insular long gyrus	1622	−33	−15	13	3
	L inferior temporal gyrus	845	−46	−56	6	4
	L calcarine sulcus	591	−10	−78	9	5
	L superior frontal gyrus - medial	543	−7	46	−1	6
	L parahippocampal gyrus	54	−32	−38	−2	A
	R supramarginal gyrus	6307	44	−37	24	1
	R middle frontal gyrus	1554	43	−8	42	2
	R calcarine sulcus	954	15	−50	14	3
	R cingulate gyrus	605	6	1	32	4
	R inferior frontal gyrus	87	28	26	−13	A
	R parahippocampal gyrus	38	26	−24	−6	B

inconsistent endings was greater when the global context was relevant than when it was irrelevant. The pattern of the bilateral cluster in the paracentral lobule and dorsal CS was also similar. The pattern differed from that of the parietal cluster, however, in that it consisted of modulations of degrees of deactivation, with more difficult integration resulting in greater deactivation (i.e., less activation).

Effect of global context relevance

This analysis identified regions involved in the process of integration sensitive to the relevance of the distal context. The parietal cluster included the inferior parietal lobule (mostly the supramarginal gyrus, SMG), superior parietal gyrus (SPG), and anterior IPS. These regions showed greater activation for endings preceded by relevant global context. In addition, the analysis based on the stricter voxel threshold found a cluster in the central insula, which showed greater activation for irrelevant global contexts.

Effect of ending local consistency

This analysis identified regions sensitive to the consistency of the ending with recent context. The analysis identified 6 clusters on the left and 4 on the right hemisphere. Laterally, fronto-parietal clusters included PreCG, which extended only dorsally on the left, posterior superior frontal gyrus (SFG; which extended more laterally on the right and more medially on the left), the middle portion of CS (only on the left), the middle part of the postcentral sulcus (PostCS) on the left, which extended more dorsally and ventrally on the right, the angular gyrus (AG; mostly on the right), and ventral SPG (only on the right). Medially, central cingulate cortex (CinG) was found bilaterally. On the left, activity also extended to the caudal part, to the cingulate sulcus, to the ventral part of the paracentral lobule, to the inferior precuneus, and to the posterior and inferior parts of SFG. One fronto-medial cluster on ventromedial prefrontal cortex (vmPFC) was present only on the left.

The posterior part of the insula (including transverse temporal gyrus; TTG) was also found bilaterally, and anterior insula only on the right. Temporal regions were also found on the left: posterior middle temporal gyrus (MTG), middle temporal sulcus, inferior temporal gyrus, with an extension to the inferior occipital gyrus. Occipital clusters included the anterior cuneus and the anterior POS bilaterally, and also extended to the lingual gyrus (LG) on the right. The analysis based on the stricter voxel threshold also found a cluster in right IFG-pars orbitalis, and bilateral parahippocampal gyrus (PHG).

For all of these clusters the patterns of neural activity were similar: Locally consistent endings were associated with greater activity than locally inconsistent endings. The activity of the locally inconsistent endings was above baseline for all clusters except in the left vmPFC and left occipital clusters, and in the right IFG and PHG clusters.

Neural activity associated with subsequent recall of story endings

This analysis identified brain regions where BOLD response during comprehension of story endings (for each of the 4 conditions separately) correlated with participants' accuracy in the cued-recall task. This task targeted inter-individual differences in the ability to recall items presented in each of the four conditions ($M = 36\%$, $sd = 20\%$). The purpose of this analysis was to identify regions associated with the cognitive processes mediating successful encoding of the endings. The results of the correlation of the recall scores with neural activity are shown in Fig. 6 and Table 3.

Neural activity associated with recall when global context was irrelevant to integration

The two conditions in which global context was irrelevant for the integration of the story endings (Irr_Cons and Irr_Inco) showed correlations between recall scores and BOLD activity in a few left hemisphere regions. Both conditions showed a negative correlation in occipital areas that included the parieto-occipital gyrus, the transverse occipital sulcus, and the superior cuneus. For the Irr_Inco condition, the negative correlation also extended to the posterior intraparietal sulcus (IPS). In addition, Irr_Inco also showed a positive correlation with middle superior temporal sulcus (STS), posterior superior temporal gyrus (STG) and posterior MTG. For the Irr_Cons condition, the analysis based on the stricter voxel threshold also found a negative correlation in the right cuneus and a positive correlation in left SFG.

Neural activity associated with recall when global context was relevant to integration

In the two conditions where global context was relevant for the story endings (Rel_Cons and Rel_Inco), recall scores were positively correlated with BOLD activity in a wider set of regions than those identified for conditions when global context was irrelevant.

Better memory for Rel_Cons endings positively correlated with activity in bilateral dorsomedial prefrontal cortex (dmPFC), anterior STS, posterior STS, MTG, and AG. In the left hemisphere the positive correlation held also for IFG-pars triangularis and the ventral part of the pars orbitalis, dorsolateral prefrontal cortex (dlPFC), and anterior PHG. On the right, positive correlation also held for anterior temporal lobes (aTL), and orbitofrontal cortex (OFC). The analysis based on the stricter voxel threshold also found positive correlations in anterior and posterior MTG and precuneus on the left, and in the middle frontal gyrus (MFG), IFG-pars triangularis, and middle occipital gyrus on the right. It also found a negative correlation in the left cuneus.

For the Rel_Inco condition, positive correlation was found in AG on the left, and posterior dmPFC on the right. Negative correlation was found bilaterally in posterior insula (larger on the left), TTG, dorsal CS, and dorsal PostCG. On the left, the negative correlation extended parietally to the dorsal PostCS and to the anterior superior parietal lobule (SPL). The analysis based on the stricter voxel threshold also identified a positive correlation in MFG and medial wall on the left, and MTG on the right.

Discussion

This research dissociates the brain regions subserving integration of linguistic information with local, proximal context from those subserving integration with global, distal context. Our investigation was

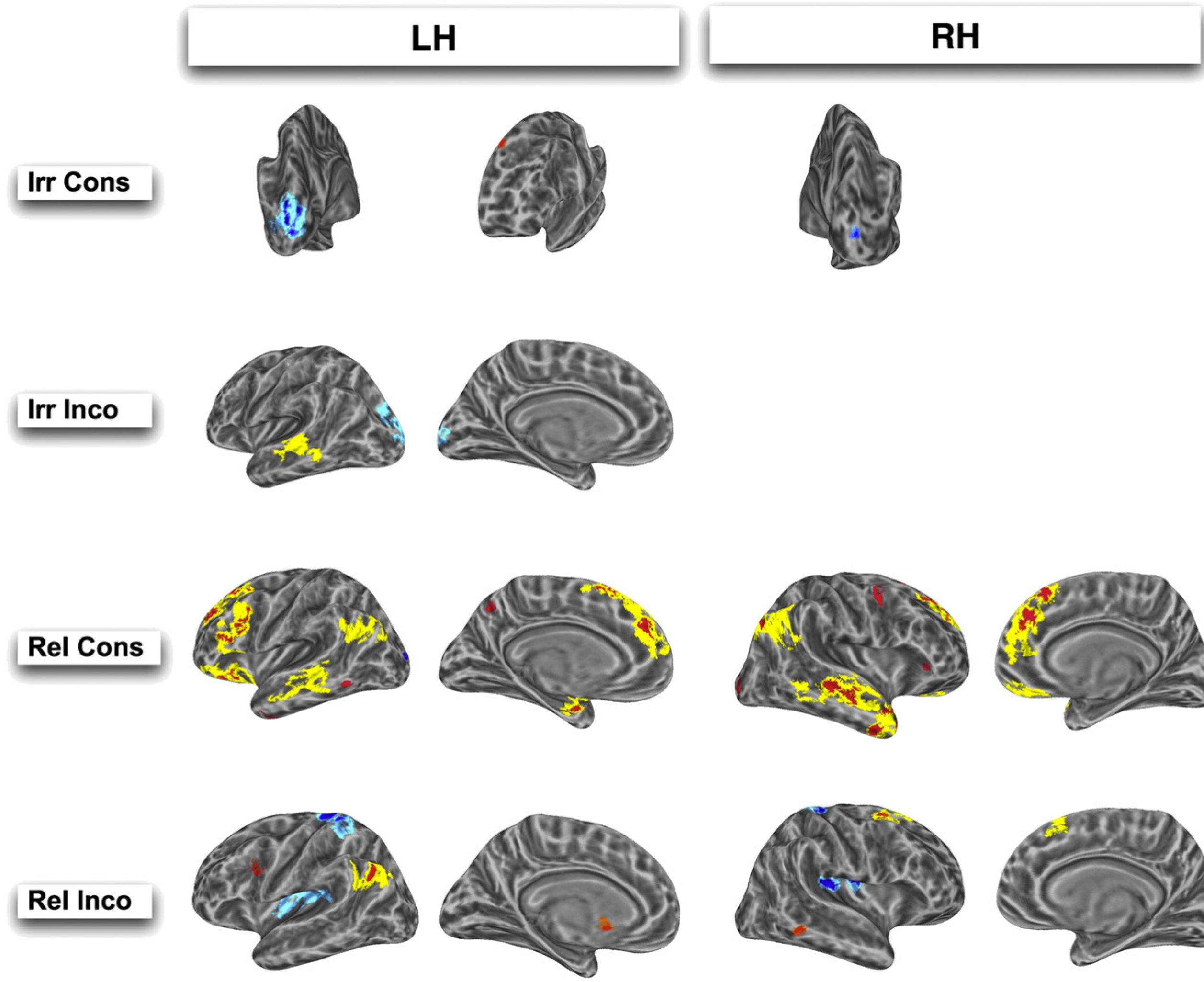

Fig. 6. Brain regions where inter-individual differences during listening to story endings correlated with recall accuracy in a post-listening cued-recall test (Experiment 2). The figure shows brain regions showing reliable correlations in each of the four experimental conditions. Yellow and light blue clusters: $p<.05$ voxel threshold. Red and dark blue clusters: $p<.005$ voxel threshold. Warm colors: positive correlations. LH = left hemisphere, RH = right hemisphere.

motivated by behavioral evidence showing that local integration is achieved differently from global integration (e.g., Egidi and Gerrig, 2006; McNamara and Magliano, 2009), thus suggesting that different underlying functional systems may be at play. This is the first study to test this hypothesis.

Our fMRI experiment identified regions associated with the processes of local and global integration, and those associated with both. This was achieved by measuring variations in the BOLD response during the comprehension of the story endings as a function of the experimental variables. Note that activity was measured while participants were listening to story endings that differed in their relation to prior context; the effects reported do not therefore indicate single-sentence semantic or syntactic processes, but integration of those sentences with prior discourse context.

Our fMRI experiment also identified brain regions associated with successful comprehension of story endings—comprehension that depends on either local or global integration. These regions are defined as those where higher success in recalling story endings correlates with BOLD responses during listening to those endings (similarly to what was done by e.g., Crinion and Price, 2005; Hasson et al., 2007).

Neural activity associated with comprehension

We had expected to identify three networks associated with comprehension as a function of the experimental variables: Two for different aspects of global integration, and one for local integration. For the global integration networks, the pattern of neural activity is consistent with our predictions that activation would be found in regions associated with top–down attention processes (for the interaction) and in regions sensitive to attention load (for the effect of global context relevance). However, for the local integration network, our expectation that the effect of local consistency of the story endings would show in regions associated with detection of linguistic inconsistencies, was not met. The results revealed instead a network indicating greater ease of processing.

With respect to the first prediction, we had hypothesized that regions showing an interaction pattern would be those previously associated with top–down attention processes, such as superior parietal regions and IPS (Cabeza et al., 2008; Corbetta and Shulman, 2002), and would thus reflect receptiveness to the most sensible story ending. The parietal regions identified by our analysis coincide with the expected network, though they additionally extended more occipitally. The pattern of signal change—with greater activity for more sensible endings—was also consistent with our interpretation of readiness. Although the additional regions found bilaterally (paracentral lobule and dorsal CS) are not usually associated with top–down attention processes, their pattern of activity suggests that in this case they support the function of the parietal cluster.

The fact that parietal activity was found in the left hemisphere is consistent with evidence suggesting the importance of the left aspect of IPS for preparatory and retrieval processes. It has been suggested

Table 3
Talairach coordinates of centers of clusters where BOLD activity during listening to story endings correlated with cued-recall performance for story endings, in each of the four experimental conditions (Experiment 2). The table also shows the direction of correlation in each cluster. In general, occipital, sensory-motor, and posterior insular regions demonstrated a negative correlation whereas temporal, frontal, and parietal areas demonstrated a positive correlation. Successful memory of endings in which global context was relevant recruited a larger set of regions. Clusters match regions in Fig. 6: Numbers refer to large clusters ($p<.05$ voxel threshold) and letters to focal clusters ($p<.005$ voxel threshold).

Condition	Region (focus point)	Area (mm^2)	x	y	z	Cluster	CorrDir
Irr Cons	L superior occipital gyrus	575	−12	−93	4	1	Neg
	L superior frontal gyrus	38	−7	33	35	A	Pos
	R right cuneus	51	7	−83	14	B	Neg
Irr Inco	L superior occipital gyrus	823	−15	−94	7	1	Neg
	L superior temporal gyrus	637	−57	−47	14	2	Pos
Rel Cons	L inferior circular sulcus of insula	1761	−38	5	−17	1	Pos
	L superior frontal gyrus - medial	1134	−9	25	28	2	Pos
	L inferior frontal sulcus	952	−35	24	19	3	Pos
	L superior temporal sulcus	753	−48	−34	9	4	Pos
	L inferior parietal gyrus	575	−36	−55	45	5	Pos
	L middle temporal gyrus	63	−51	−7	−15	A	Pos
	L precuneus	59	−9	−57	35	B	Pos
	L cuneus	56	−19	−90	5	C	Neg
	L middle temporal gyrus	45	−46	−43	1	D	Pos
	R superior temporal sulcus	1785	61	−18	−16	1	Pos
	R superior frontal gyrus - medial	1477	10	47	39	2	Pos
	R inferior parietal gyrus	795	45	−67	25	3	Pos
	R orbitofrontal gyrus	516	5	27	−15	4	Pos
	R middle frontal gyrus	102	22	36	42	A	Pos
	R cuneus	91	17	−96	5	B	Pos
	R insula	58	31	20	4	C	Pos
Rel Inco	L postcentral gyrus	1082	−18	−48	68	1	Neg
	L postcentral gyrus	991	−50	−15	16	2	Neg
	L inferior parietal gyrus	640	−35	−54	43	3	Pos
	L middle frontal gyrus	96	−39	7	39	A	Pos
	L anterior cingulate gyrus	53	−2	8	−2	B	Pos
	R superior frontal gyrus	629	28	1	46	1	Pos
	R superior temporal gyrus	584	47	−22	6	2	Neg
	R postcentral gyrus	521	24	−44	63	3	Neg
	R middle temporal gyrus	58	51	−53	−13	A	Pos

that, in preparation for a task, left IPS allows either holding information online in a continuous manner or promoting retrieval from long-term memory when needed (Phillips et al., 2009). Although this hypothesis has been proposed for task-cueing paradigms, a similar process could be at work during discourse comprehension. Behavioral results are in fact often explained with accounts of holding global context information or accessing it from long-term memory when cued (Gerrig and O'Brien, 2005; McNamara and Magliano, 2009).

These results thus extends prior language literature by pinpointing superior parietal and paracentral regions as those more likely to be associated with the top–down readiness that follows from access to distal discourse context. The regions included in these clusters seem to perform successful global assimilation of the information in a way that accounts for the behavioral results obtained in Experiment 1.

Regarding the effect of global context relevance, we had hypothesized that relevant global contexts would entail a greater attention load in working memory at the time of integration because they involve manipulation of a greater amount of information than irrelevant global contexts. We had therefore expected that dorsolateral and ventrolateral prefrontal cortices and superior parietal lobules—which are usually associated with working memory processes (Wager and Smith, 2003)—would show increased BOLD response for the integration of endings preceded by relevant global contexts (as compared to irrelevant global contexts). Our analysis revealed a parietal cluster, as expected, but no frontal cluster.

The pattern of signal change in the parietal cluster is consistent with that of prior studies (e.g., Jovicich et al., 2001) in that it shows increased activation for those contexts that entail greater attention load—the relevant global contexts. The left-lateralized location of the cluster could depend on the linguistic nature of the task (although the analysis with different regressors for the story bodies reported in Supplementary Figure B also found the involvement of parieto-occipital regions on the right). The lack of prefrontal involvement is consistent with the suggestion that while parietal areas are sensitive to encoding, maintenance, and manipulation of relevant information, prefrontal areas are recruited mostly for maintenance and manipulation (Rawley and Constantinidis, 2009; Wager and Smith, 2003). If this is the case, it is possible that in our study the difference between the relevant and irrelevant global contexts is too small to create manipulation demands that require the additional recruitment of frontal areas. (Note, however, that the analysis with different regressors for the story bodies did find a left prefrontal cluster).

A more difficult result to interpret is that found in the insula, where the increased signal change for irrelevant contexts was mostly driven by the Irr_Cons endings. The analysis with different regressors for the story bodies also found a cluster in the posterior cingulate gyrus with a similar pattern, though with different baselines. This suggests the presence of a network that distinguishes locally consistent endings preceded by irrelevant contexts as perhaps the easiest to integrate.

In general, these findings extend prior knowledge of discourse comprehension in two respects. First, by showing the brain's sensitivity to the amount of information that needs to be processed, it suggests that systems mediating discourse comprehension extend well beyond those required for sentence comprehension. Second, by highlighting the specific role of regions sensitive to the amount of distal relevant information, it allows observing how the presence of different amounts of context critically changes the way comprehension is achieved at the cortical level. Taken together, the two networks associated with the main effect of global context relevance and the interaction also further our understanding of global integration as coordinated functioning of two systems. One system accesses the distal information that is relevant at the time of integration, and the other uses it in a top–down preparation for the most sensible ending.

With respect to the effect of local ending consistency, since inconsistencies at local level ought to be easy to detect (and the behavioral data of Experiment 1 indicated that people comprehend locally consistent endings more easily), we had expected a pattern reflecting the processing costs associated with the detection of inconsistency for locally inconsistent endings. This would have been evident in greater signal change for locally inconsistent endings in a network including dmPFC, vmPFC, STS, MTG, aTL, and posterior midline regions (Ferstl et al., 2008; Hasson et al., 2007; Mason and Just, 2006). We however found that a different set of regions was sensitive to the local consistency of the endings, and that the signal change pattern was the opposite than the one expected, with increased activation for locally consistent endings.

Prior language research shows that several of the regions we found, including the left posterior insula, visual areas such as LG, cuneus, occipital gyrus, and posterior CinG, can be more strongly activated when linguistic information is easier to understand. For example, these regions are more active in bilingual speakers processing their first as compared to their second language (Perani et al., 1996, 1998; Vingerhoets et al., 2003). In addition, greater BOLD responses for sentences coherent with prior one-sentence context (as compared to incoherent ones) have been documented in posterior-central CinG, inferior precuneus, and medial PFC (Ferstl and von Cramon, 2001, 2002). Analogously, stronger activity in sensory-motor areas for locally consistent endings has been documented for sentences that are consistent with prior discourse (as compared to sentences that require an update of a story mental model; Hasson et al., 2007). This literature suggests the existence of a network that marks consistencies as the information that is more easily integrated.

As a matter of fact, our prediction that locally inconsistent endings should induce greater activation in regions associated with inconsistency resolution was based on prior literature that only loosely matches the processing demands of our experiment. Sentence comprehension studies and most discourse comprehension studies use an explicit task: Participants are asked to formulate sensibility judgments for sentences containing incongruent words or for critical incongruent sentences in stories (e.g., Ferstl and von Cramon, 2001, 2002; Ferstl et al., 2005; Kiehl et al., 2002; Kuperberg et al., 2000, 2003). Such a task is likely to induce a strategic process of inconsistency detection that may not occur, or may be performed differently, during passive listening. There is in fact evidence that passive comprehension and explicit judgments involve, at least partially, different functional networks (both at the sentence and discourse level; e.g., Chow et al., 2008; Rodd et al., 2005; Zhu et al., 2012).

Other discourse comprehension studies that have also found greater activation for incongruent than congruent language have done so by comparing coherent stories and a series of unrelated sentences (e.g., Giraud et al., 2000; Vogeley et al., 2001; Xu et al., 2005; Yarkoni et al., 2008) or coherent and ambiguous stories that only make sense when accompanied by a title or an image (as described in the Introduction; e.g., Maguire et al., 1999; Martín-Loeches et al., 2008). As sentences are rarely uttered in unrelated or ambiguous streams, these studies create extreme conditions for the language comprehension system and may not be able to capture the natural processes occurring during more ordinary circumstances of language use.

Studies that have compared the comprehension of stories with and without coherence breaks using a passive comprehension task are a closer match to our experiment. These studies have found greater activation for inconsistent than consistent information in inconsistency detection regions (Deen and McCarthy, 2010; Ferstl and von Cramon, 2001, 2002; Hasson et al., 2007), but have also found the opposite pattern in regions that we have also identified for local consistency (Ferstl and von Cramon, 2001, 2002; Hasson et al., 2007). Unlike our experiment, however, none of those studies could tease apart integration of incoming information from different amounts of prior context, and cannot fully account for the patterns of increased signal change they find for consistent information. The assumption behind all of the previously mentioned studies—and behind the predictions that we had initially formulated for the local consistency effect—is that there is only one type of integration process, and that the language systems treat all information about prior context equally. Our design, in contrast, dissociates the impact of local and global information on the integration process and shows that different processes are at play when either only local context is taken into account or global context is also included.

With respect to the local integration network that our experiment identifies, we suggest that activity in these areas indicates a process of continuous, monotonic updating of a knowledge base, which occurs when listeners construct a model of the characters' actions depicted by the endings that integrate locally. Consistent with the literature that shows increased activation in these regions for easier linguistic tasks, we therefore propose that the regions isolated by our analysis are associated with processes of fluent, continuous integration and not with processes that resolve discrepancies. Our results also suggest that local integration is an easy form of integration and therefore a strong default. The greater extension of the local integration network (as compared to the global integration networks) further supports the primacy of local over global integration.

Consistent with the results of the local integration network, the pattern of reading latencies obtained in Experiment 1 also shows comprehenders' tendency to more swiftly process locally consistent endings, with a modulation of this effect by global context. Compared to these behavioral results, the investigation of the BOLD activity constitutes a definite advantage in that it allows teasing apart different processes that may be reflected in the same pattern of reading latencies. In this case, it allows distinguishing between inconsistency detection and processing fluency, both of which produce longer reading latencies for what is detected as inconsistent and what is processed less fluently. Our data therefore suggest that the notion of consistency in theories of language should be revised to distinguish detection of logical inconsistency from processing fluency.

It is also possible that an ERP experiment could help differentiate between these processes. The N400 reflects more than one process of inconsistency detection (e.g., lexical access, integration, for a review see Kutas and Federmeier, 2011), and an ERP study using our paradigm might be able to differentiate aspects of the effect that are related to the relevance of the global context from those related to fluent, local updating. It is possible, for example, that global context relevance induces a shift in the peak of the N400. Combining fMRI and ERP results could also be useful in understanding which regions reflect different operations over time. The combination of the two methods has in fact been used successfully for a deeper understanding of the neurobiology of cognitive processes (e.g., Nieuwland, 2012; Nieuwland and Martin, 2012; St George et al., 1994, 1999).

The three networks identified by the two main effects and interaction may be part of a single system that is modulated by different cognitive processes or reflect operations in largely non-overlapping regions. To examine whether local and global consistencies are mediated by similar systems it is important to determine the overlap between the regions identified in each of the three analyses. As can be seen in Fig. 7, which shows all three effects projected on the same cortical surface, the relative overlap is extremely limited. This suggests that local and global consistencies (and their interaction) are associated with different functional systems, and are therefore processed as separate functions at the cortical level.

To summarize, this analysis reveals three different networks involved in different processes of integration. One network, which includes regions associated with top–down attention processes (superior parietal, IPS), indicates readiness to the most sensible story ending once the global context is taken into account. A second network, which includes regions sensitive to attentional load (SMG, SPL, aIPS), shows sensitivity to the amount of information considered when the global context weighs on integration. A third network, which includes many regions that do not coincide with those usually associated with inconsistency

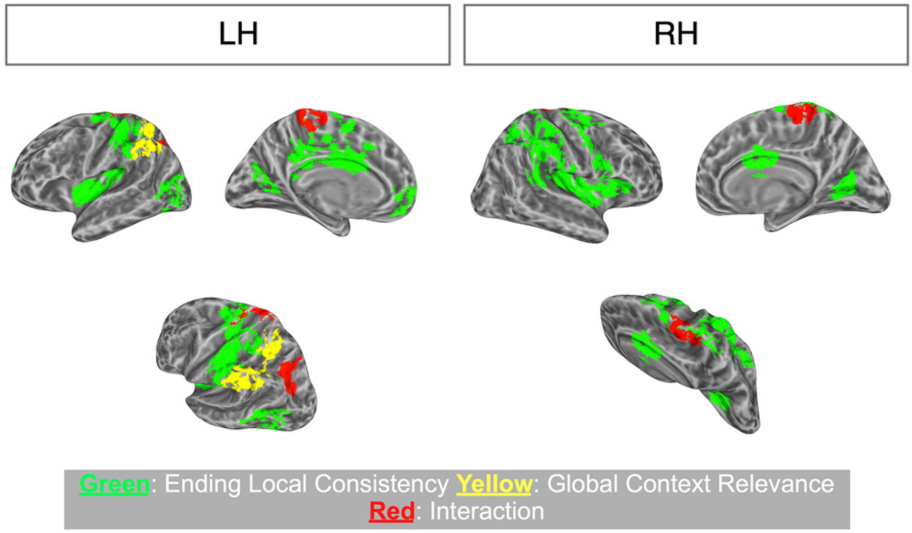

Fig. 7. Examination of region overlap between main effects and interaction in Experiment 2. Projection of the BOLD activity associated with the main effect of Ending Local Consistency, the main effect of Global Context Relevance, and their interaction. LH = left hemisphere, RH = right hemisphere.

detection, shows increased processing fluency for information that integrates locally. These are novel results in that they show how discourse integration results from the interplay of distinct networks, each sensitive to different aspects of the task. While a system monotonically adds incoming information to the most recent context, another one accesses distal relevant context, and another one, based on all available contextual information, prepares for the integration of the most sensible incoming information.

Neural activity associated with subsequent recall

In the BOLD-to-memory correlation analysis, we had expected that when global context is relevant for ending integration, subsequent memory for the endings would rely on a more extended network of brain regions—one that mediates the integration of earlier- and later-encountered information. This prediction was based on the assumption that comprehension of the ending following a relevant global context requires manipulation of a larger amount of information and makes the integration task more demanding. This was indeed the pattern found in the data, as the networks showing correlations between BOLD signal and memory scores for the two conditions in which the global context is relevant were more extensive than those identified for the two conditions in which the global context is irrelevant.

We had also expected positive correlations in a large network of prefrontal, parietal, and temporal areas, based on prior research on memory for language (Crinion and Price, 2005; Hasson et al., 2007; Maguire et al., 1999; Yarkoni et al., 2008). The pattern of results we found is in general consistent with this expectation. Importantly, however, the fact that different regions of this network show correlations in different conditions suggests that the language comprehension processes involved in each of the various conditions differ to a certain extent. Specifically, Irr_Inco shows positive correlations only in temporal areas (e.g., STG, MTG) usually associated with inconsistency detection processes both at sentence and discourse level and with linguistic semantic integration in different situations (Ferstl et al., 2008; Humphries et al., 2007). This suggests that the magnitude of the memory effect in this condition—in which the story ending was inconsistent and the global context irrelevant—depended on the system's attempt to find coherence between the inconsistent ending and prior context.

The two conditions in which the global context was relevant to integration showed correlation in a wider set of frontal (dlPFC, dmPFC), temporal (STS, MTG, aTL), and parietal (AG) regions. These regions have been linked to successful memory for language (Hasson et al., 2007; Yarkoni et al., 2008) and are known to contribute to integration of semantic information with prior context (e.g., Binder et al., 2009; Ferstl et al., 2008; Humphries et al., 2007). More specifically, AG has often been shown to be involved in processes of discourse comprehension (e.g., Menenti et al., 2009; Xu et al., 2005) and to play a role in the unification of incoming information with world knowledge (Menenti et al., 2009), ongoing verbal context (Humphries et al., 2007), and visualization of scenes derived from the discourse (Xu et al., 2005). Among the frontal areas identified by our analysis, dmPFC has been found to be involved in integration processes (e.g., Ferstl and von Cramon, 2002; Ferstl et al., 2005; Hasson et al., 2007; Xu et al., 2005). We therefore suggest that the positive correlation found in left AG and right dmPFC in the Rel_Inco condition—in which the story ending was locally inconsistent and the global context relevant—reflects a relatively smooth process of unification and integration, perhaps because the ending was globally consistent.

Among the other regions that showed positive correlations in Rel_Cons, left IFG and PHG are known to be involved in semantic processes and integration during language comprehension (e.g., Binder et al., 2009; Menenti et al., 2009; Xiang et al., 2010). OFC has been found to also contribute to the comprehension of discourse (Ferstl et al., 2005; Xu et al., 2005), although its role is not entirely clear. It has been suggested that the involvement of OFC may reflect a process of empathization with the story characters (Ferstl et al., 2005).

One additional frontal area identified by our analysis—dlPFC—has been found to contribute to the processing of semantic information in sentence comprehension (in particular on the left; Manenti et al., 2008), but it has also been shown to increase with working memory demands (e.g., Wager and Smith, 2003). In Rel_Cons, the positive correlation between dlPFC activity and accurate memory may reflect the increased working memory load required by the integration of the endings when a greater global context is relevant and is held available at the time of integration.

The positive correlations between BOLD activity and memory in the Rel_Cons condition also mark the major extent of the default-mode network (Raichle et al., 2001). A speculative function of this network is that of mental simulation of hypothetical scenarios (Buckner et al., 2008). It is likely that, for the successful integration of the endings with the rest of the story, the network provides the imaginative construction of the events depicted by the story endings and their narrative context.

Thus, these results suggest that when global contexts are relevant for the comprehension of the endings, successful memory for endings consistent only with local context depends on the activity of a large language network, especially when the endings are locally consistent (and globally inconsistent). This network detects the inconsistency of the endings with global context, performs semantic integration with prior context, is sensitive to the increased working memory load of the global context, engages in some imaginative process of the events depicted in the endings, and possibly allows some degree of empathization with the story characters.

The analysis of the neural activity associated with recall also documented unexpected negative correlations in all conditions except for Rel_Cons. The negative correlation in occipital areas found in Irr_Cons and Irr_Inco is consistent with recent work showing that visual cortex activity can track auditory demands (Hertz and Amedi, 2010) and that cross-modal deactivation of the visual areas during auditory discrimination correlates positively with task difficulty (Hairston et al., 2008). This suggests that more successful memory for consistent auditory discourse entails reduced visual processing. Finally, the negative correlation in dorsal motor and sensory regions found in Rel_Inco is consistent with prior literature showing that increased activity in these regions is associated with decreased comprehension of action-related sentences (Beilock et al., 2008). In our results, the negative correlation between memory and sensory-motor activity was also accompanied by a negative correlation with posterior insula, a region responsible for primary auditory functions, somatosensory processing, and interoceptive awareness (Augustine, 1996; Craig, 2002, 2009; Karnath et al., 2005).

Although some of the results mentioned above have been interpreted as a suppression of baseline neural activity within an otherwise inactive region (e.g., Hairston et al., 2008), we cannot argue for the same mechanism in this case. As can be seen in Table 4, which shows the percent of voxels demonstrating positive and negative activity in our group of participants, there is no conclusive evidence of suppression—many of the voxels in the regions showing negative correlation in fact showed above baseline activity. These results therefore suggest that successful memory for endings requires attenuation, but not necessarily inhibition, in processing visual information, bodily movements, sensations, and self-awareness. This attenuation perhaps occurs to allow greater focus on the linguistic stimuli and facilitate the functioning of the regions responsible for the actual integration.

Table 4
Breakup of voxel signs in regions showing correlations with memory. The figure shows the percentage of voxels with Betas reliably above and below zero for each cluster showing correlation with memory scores, for at least 11 of 14 total participants (p<.05 on sign test). Negative correlations are highlighted in gray. Even clusters showing negative correlations have more positive than negative activations, indicating that these clusters do not show deactivation.

		LH		RH	
	Cluster	Pos	Neg	Pos	Neg
Irr Cons		42	0		
Irr Inco	1	25	4		
	2	71	0		
Rel Cons	1	44	0	91	0
	2	61	0	70	0
	3	77	0	77	0
	4	81	0	14	0
	5	48	1		
Rel Inco	1	1	14	38	0
	2	3	0	2	0
	3	49	0	0	16

In conclusion, this analysis yielded three crucial results. First, when the global context is relevant for the integration of story endings, successful memory for the endings depends on greater activity within a larger brain network. This is in keeping with the pattern of recall scores, which shows better memory for endings preceded by relevant global contexts. Second, the pattern of positive correlations shows that better memory is associated with greater activation in regions usually involved in linguistic processing and memory for language. This result therefore provides additional evidence of the contribution of these regions to language integration.

Finally, the pattern of negative correlations in visual, sensory, and motor areas is a novel result in discourse comprehension (although it has a long history in other domains; e.g., Drevets et al., 1995). It suggests that successful memory requires decreased visual processing, bodily movements and awareness, perhaps to facilitate the functioning of the regions responsible for the actual integration by decreasing resource allocation to irrelevant tasks. In general, these results suggest that successful integration requires the recruitment not only of regions associated with active assimilation, but also of regions that keep under control potentially distracting information from other sources.

Conclusion

Taken together, our results extend our understanding of the neurobiological substrates of discourse comprehension (e.g., Ferstl et al., 2008; Mason and Just, 2006) by identifying networks associated with three important aspects of local and global integration: receptiveness to endings that make most sense both locally and globally, sensitivity to the availability of global context, and greater processing fluency for information that is consistent with the local context. They also extend our knowledge of integration processes in showing a decreased deployment of resources for tasks irrelevant to comprehension.

In addition, these results demonstrate how the integration of ongoing discourse is carried out in relation to the local and global contexts. While some systems are sensitive to the holistic relation of content to both the local and global contexts, other regions play a more generic role, tracking solely the relevance of the global context, or the local consistency of an ending, or the control over potentially distracting information. It is likely that the interaction among these systems during processing is key for determining successful comprehension, since a failure of retrieval of the global context, a failure to keep distracting information at bay, or an exaggerated emphasis on the local context can result in miscomprehension of discourse as a whole. Successful comprehension is only possible through the cooperation of different brain systems each recruited for different aspects of the process.

Acknowledgments

We thank Alessia Giovenzana and Ketti Mazzocco for their contribution to the preparation of the stimuli and Silvia Modenese for her help in coding the recall protocols. This research was supported in part by the Fondazione Cassa di Risparmio di Trento e Rovereto.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2013.01.003.

References

Albrecht, J.E., O'Brien, E.J., 1993. Updating a mental model: maintaining both local and global coherence. J. Exp. Psychol. Learn. Mem. Cogn. 19, 1061–1070.
Augustine, J.R., 1996. Circuitry and functional aspects of the insular lobe in primates including humans. Brain Res. Rev. 22, 229–244.

Beilock, S.L., Lyons, I.M., Mattarella-Micke, A., Nusbaum, H.C., Small, S.L., 2008. Sports experience changes the neural processing of action language. Proc. Natl. Acad. Sci. U. S. A. 105, 13269–13273.

Binder, J.R., Desai, R.h, Graves, W.W., Conant, L.L., 2009. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cereb. Cortex 19, 2767–2796.

Bransford, J.D., Johnson, M.K., 1972. Contextual prerequisites for understanding: some investigations of comprehension and recall1. J. Verbal Learn. Verbal Behav. 11, 717–726.

Buckner, R.L., Andrews-Hanna, J.R., Schacter, D.L., 2008. The brain's default network. Ann. N. Y. Acad. Sci. 1124, 1–38.

Cabeza, R., Ciaramelli, E., Olson, I.R., Moscovitch, M., 2008. The parietal cortex and episodic memory: an attentional account. Nat. Rev. Neurosci. 9, 613–625.

Chow, H.M., Kaup, B., Raabe, M., Greenlee, M.W., 2008. Evidence of fronto-temporal interactions for strategic inference processes during language comprehension. NeuroImage 40, 940–954.

Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. Nat. Rev. Neurosci. 3, 201–215.

Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput. Biomed. Res. 29, 162–173.

Craig, A.D., 2002. How do you feel? Interoception: the sense of the physiological condition of the body. Nat. Rev. Neurosci. 3, 655–666.

Craig, A., 2009. How do you feel—now? The anterior insula and human awareness. Nat. Rev. Neurosci. 10, 59–70.

Crinion, J., Price, C.J., 2005. Right anterior superior temporal activation predicts auditory sentence comprehension following aphasic stroke. Brain 128, 2858–2871.

Crinion, J.T., Lambon-Ralph, M.A., Warburton, E.A., Howard, D., Wise, R.J.S., 2003. Temporal lobe regions engaged during normal speech comprehension. Brain 126, 1193–1201.

Deen, B., McCarthy, G., 2010. Reading about the actions of others: biological motion imagery and action congruency influence brain activity. Neuropsychologia 48, 1607–1615.

Drevets, W.C., Burton, H., Videen, T.O., Snyder, A.Z., Simpson, J.R., Raichle, M.E., 1995. Blood flow changes in human somatosensory cortex during anticipated stimulation. Nature 373, 249–252.

Egidi, G., Gerrig, R.J., 2006. Readers' experiences of characters' goals and actions. J. Exp. Psychol. Learn. Mem. Cogn. 32, 1322–1329.

Egidi, G., Nusbaum, H.C., 2012. Emotional language processing: how mood affects integration processes during discourse comprehension. Brain Lang. 122, 199–210.

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., von Cramon, D.Y., 2008. The extended language network: a meta-analysis of neuroimaging studies on text comprehension. Hum. Brain Mapp. 29, 581–593.

Fischl, B., Sereno, M.I., Dale, A.M., 1999. Cortical surface-based analysis II: inflation, flattening, and a surface-based coordinate system. NeuroImage 9, 195–207.

Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A., Noll, D.C., 1995. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. Magn. Reson. Med. 33, 636–647.

Gerrig, R.J., O'Brien, E.J., 2005. The scope of memory-based processing. Discourse Process 39, 225–242.

Giraud, A.L., Truy, E., Frackowiak, R.S.J., Grégoire, M.C., Pujol, J.F., Collet, L., 2000. Differential recruitment of the speech processing system in healthy subjects and rehabilitated cochlear implant patients. Brain 123, 1391–1402.

Graesser, A.C., Singer, M., Trabasso, T., 1994. Constructing inferences during narrative text comprehension. Psychol. Rev. 101, 371–395.

Guéraud, S., Harmon, M.E., Peracchi, K.A., 2005. Updating situation models: the memory-based contribution. Discourse Process 39, 243–263.

Hagoort, P., Van Berkum, J., 2007. Beyond the sentence given. Phil. Trans. R. Soc. Biol. Sci. 362, 801–811.

Hairston, W.D., Hodges, D.A., Casanova, R., Hayasaka, S., Kraft, R., Maldjian, J.A., Burdette, J.H., 2008. Closing the mind's eye: deactivation of visual cortex related to auditory task difficulty. Neuroreport 19, 151–154.

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, 2899–2913.

Hertz, U., Amedi, A., 2010. Disentangling unisensory and multisensory components in audiovisual integration using a novel multifrequency fMRI spectral analysis. NeuroImage 52, 617–632.

Humphries, C., Binder, J.R., Medler, D.A., Liebenthal, E., 2007. Time course of semantic processes during sentence comprehension: an fMRI study. NeuroImage 36, 924–932.

Jovicich, J., Peters, R.J., Koch, C., Braun, J., Chang, L., Ernst, T., 2001. Brain areas specific for attentional load in a motion-tracking task. J. Cogn. Neurosci. 13, 1048–1058.

Kansaku, K., Yamaura, A., Kitazawa, S., 2000. Sex differences in lateralization revealed in the posterior language areas. Cereb. Cortex 10, 866–872.

Karnath, H.O., Baier, B., Nägele, T., 2005. Awareness of the functioning of one's own limbs mediated by the insular cortex? J. Neurosci. 25, 7134–7138.

Kiehl, K.A., Laurens, K.R., Liddle, P.F., 2002. Reading anomalous sentences: an event-related fMRI study of semantic processing. NeuroImage 17, 842–850.

Kintsch, W., 1988. The role of knowledge in discourse comprehension: a construction-integration model. Psychol. Rev. 95, 163–182.

Kintsch, W., 1994. Text comprehension, memory, and learning. Am. Psychol. 49, 294–303.

Kintsch, W., 1998. Comprehension: A Paradigm for Cognition. Cambridge Univ Press, New York.

Kuperberg, G.R., McGuire, P., Bullmore, E., Brammer, M., Rabe-Hesketh, S., Wright, I., David, A., 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., Holcomb, P.J., Sitnikova, T., Greve, D., Dale, A.M., Caplan, D., 2003. Distinct patterns of neural modulation during the processing of conceptual and syntactic anomalies. J. Cogn. Neurosci. 15, 272–293.

Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 62, 621–647.

Loftus, G.R., Masson, M.E., 1994. Using confidence intervals in within-subject designs. Psychon. Bull. Rev. 1, 476–490.

Long, D.L., Lea, R.B., 2005. Have we been searching for meaning in all the wrong places? Defining the "search after meaning" principle in comprehension. Discourse Process 39, 279–298.

Magliano, J.P., Radvansky, G.A., 2001. Goal coordination in narrative comprehension. Psychon. Bull. Rev. 8, 372–376.

Maguire, E.A., Frith, C.D., Morris, R., 1999. The functional neuroanatomy of comprehension and memory: the importance of prior knowledge. Brain 122, 1839–1850.

Manenti, R., Cappa, S., Rossini, P.M., Miniussi, C., 2008. The role of the prefrontal cortex in sentence comprehension: An rTMS study. Cortex 44, 337–344.

Martín-Loeches, M., Casado, P., Hernández-Tamames, J.A., Álvarez-Linera, J., 2008. Brain activation in discourse comprehension: a 3 t fMRI study. NeuroImage 41, 614–622.

Mason, R.A., Just, M.A., 2006. Neuroimaging contributions to the understanding of discourse processes, In: Traxler, M., Gernsbacher, M.A. (Eds.), Handbook of Psycholinguistics, Second ed. pp. 765–799.

McKoon, G., Ratcliff, R., 1992. Inference during reading. Psychol. Rev. 99, 440–466.

McNamara, D.S., Kintsch, W., 1996. Learning from texts: effects of prior knowledge and text coherence. Discourse Process 22, 247–288.

McNamara, D.S., Magliano, J., 2009. Toward a comprehensive model of comprehension. Psychol. Learn. Motiv. 51, 297–384.

McNamara, D.S., Kintsch, E., Songer, N.B., Kintsch, W., 1996. Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. Cogn. Instr. 14, 1–43.

Menenti, L., Petersson, K.M., Scheeringa, R., Hagoort, P., 2009. When elephants fly: differential sensitivity of right and left inferior frontal gyri to discourse and world knowledge. J. Cogn. Neurosci. 21, 2358–2368.

Nieuwland, M.S., 2012. Establishing propositional truth-value in counterfactual and real-world contexts during sentence comprehension: differential sensitivity of the left and right inferior frontal gyri. NeuroImage 59, 3433–3440.

Nieuwland, M.S., Martin, A.E., 2012. If the real world were irrelevant, so to speak: the role of propositional truth-value in counterfactual sentence comprehension. Cognition 122, 102–109.

Nieuwland, M.S., Van Berkum, J.J., 2006. When peanuts fall in love: N400 evidence for the power of discourse. J. Cogn. Neurosci. 18, 1098–1111.

Papathanassiou, D., Etard, O., Mellet, E., Zago, L., Mazoyer, B., Tzourio-Mazoyer, N., 2000. A common language network for comprehension and production: a contribution to the definition of language epicenters with PET. NeuroImage 11, 347–357.

Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, S.F., Dupoux, E., Fazio, F., Mehler, J., 1996. Brain processing of native and foreign languages. Neuroreport 7, 2439–2444.

Perani, D., Paulesu, E., Galles, N.S., Dupoux, E., Dehaene, S., Bettinardi, V., Cappa, S.F., Fazio, F., Mehler, J., 1998. The bilingual brain. Proficiency and age of acquisition of the second language. Brain 121, 1841.

Phillips, J.S., Velanova, K., Wolk, D.A., Wheeler, M.E., 2009. Left posterior parietal cortex participates in both task preparation and episodic retrieval. NeuroImage 46, 1209–1221.

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.

Rawley, J.B., Constantinidis, C., 2009. Neural correlates of learning and working memory in the primate posterior parietal cortex. Neurobiol. Learn. Mem. 91, 129–138.

Rodd, J.M., Davis, M.H., Johnsrude, I.S., 2005. The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. Cereb. Cortex 15, 1261–1269.

Saad, Z.S., Glen, D.R., Chen, G., Beauchamp, M.S., Desai, R., Cox, R.W., 2009. A new method for improving functional-to-structural MRI alignment using local Pearson correlation. NeuroImage 44, 839–848.

St George, M., Mannes, S., Hoffman, J.E., 1994. Global semantic expectancy and language comprehension. J. Cogn. Neurosci. 6, 70–83.

St George, M., Kutas, M., Martinez, A., Sereno, M., 1999. Semantic integration in reading: engagement of the right hemisphere during discourse processing. Brain 122, 1317–1325.

van Berkum, J.J.A., Hagoort, P., Brown, C.M., 1999. Semantic integration in sentences and discourse: evidence from the N400. J. Cogn. Neurosci. 11, 657–671.

van Berkum, J.J.A., Zwitserlood, P., Hagoort, P., Brown, C.M., 2003. When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. Cogn. Brain Res. 17, 701–718.

Vingerhoets, G., van Borsel, J., Tesink, C., van den Noort, M., Deblaere, K., Seurinck, R., Vandemaele, P., Achten, E., 2003. Multilingualism: an fMRI study. NeuroImage 20, 2181–2196.

Vogeley, K., Bussfeld, P., Newen, A., Herrmann, S., Happe, F., Falkai, P., 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. Cogn. Affect. Behav. Neurosci. 3, 255–274.

Xiang, H.D., Fonteijn, H.M., Norris, D.G., Hagoort, P., 2010. Topographical functional connectivity pattern in the perisylvian language networks. Cereb. Cortex 20, 549–560.
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., Zacks, J.M., 2008. Neural substrates of narrative comprehension and memory. NeuroImage 41, 1408–1425.
Zaitsev, M., Hennig, J., Speck, O., 2004. Point spread function mapping with parallel imaging techniques and high acceleration factors: fast, robust, and flexible method for echo-planar imaging distortion correction. Magn. Reson. Med. 52, 1156–1166.
Zhu, Z., Hagoort, P., Zhang, J.X., Feng, G., Chen, H.-C., Bastiaansen, M., Wang, S., 2012. The anterior left inferior frontal gyrus contributes to semantic unification. NeuroImage 60, 2230–2237.
Zwaan, R.A., Radvansky, G.A., 1998. Situation models in language comprehension and memory. Psychol. Bull. 123, 162–185.
Zwaan, R.A., Langston, M.C., Graesser, A.C., 1995. The construction of situation models in narrative comprehension. Psychol. Sci. 6, 292–297.