Published in final edited form as:

Neuroimage. 2008 April 1; 40(2): 932–939.

NARRATIVE SPEECH PRODUCTION: AN fMRI STUDY USING CONTINUOUS ARTERIAL SPIN LABELING

Vanessa Troiani, Maria A. Fernández-Seara, Ze Wang, John A. Detre, Sherry Ash, and Murray Grossman

Department of Neurology, University of Pennsylvania School of Medicine

Abstract

Functional magnetic resonance imaging (fMRI) with continuous arterial spin labeling (CASL) was employed to monitor brain activation during narrative production of a semi-structured speech sample in healthy young adults. Subjects were asked to describe a wordless children's picture story. Significant activations were found in bilateral prefrontal and left temporal-parietal regions during narrative production relative to description of a single picture and relative to viewing the wordless picture story while producing a nonsense word. We conclude that an inferior frontal cortex serves as a top-down organizational resource for narrative comprehension and demonstrate the feasibility of collecting extended speech samples using CASL perfusion fMRI.

Keywords

narrative; speech; fMRI; prefrontal; arterial spin labeling

INTRODUCTION

Narrative production involves organizing and expressing a complex series of events. This process is fundamental for human communication, yet we know little about its neural basis. Our model of narrative production involves at least two components, including a linguistic component and an executive resource component (Mar, 2004). Linguistic functions implicated during narrative production include phonology, morphology, lexical, and grammatical processing, which serve to express the content of a story. The focus of this study is the second component, involving higher level cognitive processes that play a crucial role in organizing a narrative, such as sustaining a story's theme through working memory, and maintaining story coherence through top-down planning and organization. These linguistic and cognitive processes must successfully interact to produce a sequence of utterances that relate to each other in expressing a logical and coherent narrative. In this report, we use functional magnetic resonance imaging (fMRI) with continuous arterial spin labeling (CASL) to monitor regional brain activation during narrative speech production.

Please address correspondence to: Murray Grossman, Department of Neurology – 3 West Gates, University of Pennsylvania School of Medicine, 3400 Spruce St, Philadelphia, PA 19104-4283, Voice: 215-349-5863; fax: 215-349-8464; email: mgrossma@mail.med.upenn.edu.

Portions of this work were presented at the American Academy of Neurology Annual Meeting, Boston, 2007.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Observations of non-aphasic patients with focal central nervous system damage implicate frontal cortex in the higher-level processes contributing to narrative. For example, patients with executive dysfunction due to prefrontal damage following traumatic brain injury fail to construct cohesive, temporally sequenced speech samples (Ferstl, Guthke, & von Cramon, 1999; Ferstl & von Cramon, 2002). This dysfunction can appear after damage to left or right prefrontal cortex. Patients with Alzheimer's disease also have narrative deficits (Chapman et al., 2002; Glosser & Deser, 1992). Deficits in narrative production in AD have also been attributed to executive dysfunction, although their episodic memory difficulty makes it difficult to determine the basis for their poor narrative performance.

Perhaps the most informative studies come from non-aphasic patients with frontotemporal dementia (FTD) who have social and executive limitations but minor memory impairments (Libon et al., 2007). These patients experience significant narrative difficulty (Ash et al., 2006; Chapman et al., 2005). A detailed analysis of performance when narrating a wordless picture story shows these patients have a limited grasp of the story's overall theme and poor connectedness between specific events in their stories, even though lexical and grammatical aspects of word and sentence use are relatively preserved (Ash et al., 2006). Additional evidence consistent with impaired top-down organization comes from a selective deficit at detecting errors in the ordering of script events in these social and executive FTD patients, and their greater deficit with event ordering than the detection of semantic substitutions in a script (Cosentino, Chute, Libon, Moore, & Grossman, 2006). FTD patients with social and executive deficits tend to have cortical atrophy in prefrontal, ventral frontal, and anterior temporal brain regions, often more prominently on the right than the left (Grossman et al., 2004; Rosen et al., 2005; Williams, Nestor, & Hodges, 2005). Direct evidence relating narrative limitations to a specific neuroanatomic substrate comes from a correlative study of these non-aphasic FTD patients. Using a voxel-based morphometric analysis of cortical volume, Ash et al (2006) observed a specific correlation between limited connectedness in story narrative and cortical atrophy in prefrontal, inferior frontal, and temporal regions of FTD patients.

In the current study, we sought converging evidence about the neural basis for narrative discourse in healthy young adults from fMRI. Several studies have used fMRI to monitor the appreciation of narrative coherence, finding that bilateral medial and lateral frontal and anterior temporal regions are activated while listening to a narrative relative to baselines involving rest or unrelated sentences (Fletcher et al., 1995; Gallagher, 2000; Mazoyer et al., 1993; Xu, Kemeny, Park, Frattali, & Braun, 2005). A more explicit approach to the evaluation of narrative reported left medial prefrontal activation during explicit judgments of the coherence of pairs of sentences (Ferstl et al., 2002) or coherence judgments depending on the presence of definite articles (Robertson et al., 2000). In two recent studies, subjects were asked to judge the relatedness of a third sentence to two prior sentences establishing the scene (Kuperberg, Lakshmanan, Caplan, & Holcomb, 2006; Mason & Just, 2004). This work reported bilateral frontal, temporal, and parietal activation that was greatest for stimuli demanding a coherent linkage between sentences.

We were particularly interested in monitoring regional brain activity during a more natural and ecologically valid task than passively listening to narrative – narrative speech production. One strategy for studying narrative production involves examining activation during inner speech (Wildgruber, Ackermann, Klose, Kardatzki, & Grodd, 1996; Wildgruber, Ackermann, & Grodd, 2001). The results of studies involving inner speech are often compatible with those involving actual speech production (Palmer et al., 2001; Rosen, Ojemann, Ollinger, & Petersen, 2000; Shuster & Lemieux, 2005), although this work has focused on single words. Moreover, we wanted to ensure that we were monitoring brain activity during coherent, full narrative production where discourse organization must be explicit rather than during a kind of mental shorthand where organizational links crucial to the coherence of a narrative may not be fully

realized. We also sought to minimize the methodological confound of the absence of auditory self-stimulation during the production of silent inner speech, and we wanted to be able to monitor narrative accuracy (Munhall, 2001).

Monitoring brain activity during speech requires us to limit motion artifact. In one influential study looking at narrative production, activity was monitored during an extended speech sample with PET (Braun, Guillemin, Hosey, & Varga, 2001). A broad range of brain regions were activated during subjects' extemporaneous accounts of personal experiences. Among these were inferior, dorsolateral, and medial frontal regions as well as lateral temporal regions, more prominently on the left than the right, and bilateral temporal-occipital regions. Since narratives were produced with both oral speech and gestural sign language in this study, activations seen in this conjunction analysis were common to both modalities and thus could not be easily attributed to modality-specific components such as a motor speech mechanism. This study is extraordinarily important because of its ecological validity, but the target of the narrative was not known and thus the ability to validate the content of the narrative was limited.

Additional reports describe studies of speech production during fMRI (Palmer et al., 2001; Rosen et al., 2000; Shuster et al., 2005). fMRI has many advantages over PET due to its superior spatial resolution, non-invasiveness, and absence of ionizing radiation. However, many challenges must be overcome during the collection of speech data in the bore of a magnet (Birn, Cox, & Bandettini, 2004; Xu et al., 2005). These are due to facial muscle and tongue use, head movement, and other artifact-producing aspects of speech production. Moreover, we sought to collect data over an extended period of time, that is, in a manner that more closely resembles the extended narrative discourse that we use in our day-to-day speech. This was achieved by asking participants to formulate a narrative consisting of an extended speech sample based on a large sequence of pictures illustrating a children's picture story. We circumvented potential problems associated with fMRI data collection during speech by using sparse data sampling, taking advantage of the delay in the hemodynamic response to collect cortical activation data following the cessation of speech (Belin et al., 1998; Hall et al., 1999). We also elected to use perfusion fMRI based on an arterial spin labeling (ASL) perfusion technique rather than Blood Oxygen-Level Dependent (BOLD) fMRI because ASL contrast does not depend on susceptibility effects, and ASL perfusion fMRI can be obtained using imaging sequences that minimize artifacts from static susceptibility gradients (Fernandez-Seara et al., 2005; Kemeny, Ye, Birn, & Braun, 2005; Kim et al., 2006). Further, because labeled and control images are interleaved during ASL scanning, ASL perfusion fMRI data do not suffer from the low frequency noise that is typically present in BOLD data, and sustained task conditions can be studied without any penalty (Wang et al., 2003).

Finally, we attempted to devise baseline conditions that would allow us to separate the top-down organization component of extended narrative apart from factors like the meaningfulness of the language being produced, the gist level of narrative appreciation associated with an approximation of story organization, and the stimulation associated with hearing one's own voice. Our first baseline involved the description of single pictures that could not be assembled into an extended narrative. This allowed us to control for grammatical, phonologic, and lexical semantic content of speech while minimizing the narrative component involved in describing the story associated with a long series of pictures. The second baseline involved looking at the ordered pictures of the entire story while producing a nonsense word to control for low level motor components of articulation and suppress executive resources associated with inner speech such as the phonologic loop of working memory. Based on our previous experience with correlative studies of narrative speech deficits with frontal cortical atrophy in FTD (Ash et al., 2006) and previous studies of narrative speech in a PET environment (Braun et al., 2001), we hypothesized that we would see activations consistent with our two-component

model of narrative, including bilateral frontal activation to support narrative organization and temporal-parietal activation to support story content during narrative expression.

METHODS

Subjects

Participants included 15 healthy, right-handed, native English speakers (10 females and 5 males, mean age=24.4 years; mean education=15.8 years) from the University of Pennsylvania community who were paid for their participation. They completed an informed consent procedure approved by the Institutional Review Board at the University of Pennsylvania.

Materials

We elicited a semi-structured speech sample from the subjects with the wordless children's picture book *Frog, Where Are You* (Mayer, 1969). The story opens with a boy and his dog admiring a pet frog in a large jar as they prepare to go to sleep. During the night, the frog climbs out of the jar and escapes through the window. The boy and the dog have several adventures in the nearby woods as they search for the frog. They encounter a groundhog, a hive of bees, an owl, and a deer. They finally find their frog, only to discover that he is with a lady frog, and they have a group of baby frogs. The boy and dog return home with one of the babies, waving a cheerful goodbye to the frog family.

We studied narrative discourse using this target story rather than free conversation because the pictures allowed us to determine the accuracy of production relative to the intended target. We did not ask participants to tell us an over-learned story such as a fairy tale because the automatic nature and varying content of such stories potentially confounded our ability to examine narrative organization. We also did not ask subjects to describe a single scene such as the Cookie Theft picture because the material in a single scene is not complex enough to evoke a rich narrative.

Procedure

The task in the magnet consisted of 5 blocks, each made up of the 24 black-and-white drawings of the target story and each lasting 4.2 minutes. The blocks were presented in a specific order designed to minimize the subject's ability to infer the story narrative while collecting imaging data corresponding to baseline conditions in a within-subject design. Each stimulus picture in a block was presented during a 9000 ms sequence (Figure 1). Subjects were presented with a 100 ms cross (+) which served as a warning cue, followed by an image from the story for 2600 ms. At this point, a green screen flashed for 200 ms, indicating that the subject should begin speaking. The same image was then presented for another 5300 ms, and subjects were cued to stop speaking with a red screen for 400 ms prior to a 400 ms period of brain volume acquisition, during which a blank screen was viewed. Instructions were given aurally to the subject prior to each block using MRI-compatible headphones with a noise cancellation system (Avotech), and instructions were also presented visually on the screen for the first 36 seconds of each block. The data corresponding to this period were discarded during subsequent imaging analysis.

The first block involved the presentation of pictures in a fixed random order and the simultaneous production of a repeated pseudoword ("padaka"). Pre-testing in young controls confirmed that a story could not be inferred from the presented order of pictures. The second block involved the presentation of randomly-ordered pictures (in a different order than the previous block, and one where the story could not be inferred) while subjects named specific objects indicated on the picture. The study reported below focused on the following three blocks:

Block 3 (randomly-ordered pictures and single picture description)—In this block, subjects briefly described each picture presented in a fixed random order differing from the previous two orderings (again, the story could not be inferred from this ordering). Instructions were: *Following a “+” cue, a picture will be presented. When the green screen flashes, you may begin speaking. Please briefly describe the picture with a sentence or two. Please stop when the red screen flashes.*

Block 4 (story-order pictures and pseudoword production)—The drawings were presented in the order of the story. This was a subject’s first knowledge that the pictures constituted a story and that the experiment involved a story. During presentation of each picture, subjects repeatedly produced a pseudoword (“padaka”). Instructions were: *You will now be shown the pictures in the order of a story. Please pay attention to the ongoing story, as you will be asked to narrate it in the future. Following a “+” cue, a picture will be presented. When the green screen flashes, you may begin speaking. Please repeat the word ‘padaka’ when viewing the picture. Please stop when the red screen flashes.*

Block 5 (story-order pictures and story description)—The drawings were presented in the order of the story, and subjects described the story. Instructions were: *You will now narrate the story. Following a “+” cue, a picture will be presented. When the green screen flashes, you may begin speaking. Please tell the story as if you are telling a child. Please stop when the red screen flashes.* We were able to record speech samples from 13 individuals while in the scanner (2 were not acquired due to a technical error during recording of the samples).

Subjects were trained ahead of time about the meaning of various cues, and the onset and offset of speech was practiced prior to entering the scanner for the study. Prior to each block, subjects were given a brief practice session. Drawings for the practice session were from a different book by the same artist as those used for the test stimuli. For the practice, Block 3 contained 5 drawings, and Block 4 and 5 contained 10 drawings. Subjects practiced beginning to speak after the green screen stimulus and stopping their speech immediately when they saw the red screen stimulus. Speech was produced and recorded during a silent period in the scanning sequence. Images were acquired while the subject was not speaking, 400 ms after the speech production for that picture stimulus. Subject responses were recorded using a microphone (Avotech) attached to headphones, designed for MR use.

Narrative Analysis

The recordings of the narratives were transcribed in detail by trained transcribers using Praat software (Institute of Phonetic Sciences, University of Amsterdam, Amsterdam, The Netherlands, available from www.praat.org). The narratives were independently scored from the transcripts by two trained judges. The narratives were analyzed for features of phonological complexity, fluency, complexity of grammatical structures, and narrative coherence. The following variables were coded:

Syllables—Total number of syllables was counted for each picture presentation as a measure of phonological complexity.

Total words produced—Every complete word in the narrative was counted, including repetitions.

Syntax—Sentence complexity was coded using an ordinal scale. An ungrammatical sentence was scored “zero.” A well-formed, grammatically correct sentence with a simple active declarative structure was scored “one.” A complex sentence, including either a dependent clause or a phrasal adjunct, was scored “two.” A compound sentence, with at least two

independent clauses, was scored “three.” A sentence containing an embedded clause was scored “four.”

Narrative Element—Sentences were given a point if the sentence referred to an element not present in the current picture. Because two subjects’ narratives were not obtained due to a recording error, the reported statistics are based on 13 narratives.

MRI Acquisition and Analysis

The experiment was carried out on a 3T Siemens Trio Scanner (Siemens Medical Systems, Erlangen, Germany). A bird-cage head coil (Bruker) was used for signal transmission and reception. Each imaging session began with a 3D MPRAGE sequence (TR=1620 msec, TE=3.9 msec, TI=950 msec, $192 \times 256 \times 160$ matrix, flip angle = 15°) which acquired anatomical images with 1 mm isotropic resolution. We used a continuous ASL (CASL) technique to monitor brain activity during narrative production tasks. CASL involves labeling the blood water as an endogenous tracer as it passes through the carotid arteries in the neck, similar to the administration of radio-labeled water in a PET study. Perfusion can be monitored continuously in a reliable manner that is independent of factors such as scanner drift during a study by the alternate administration of labeled and control series throughout the study (Aguirre, Detre, Zarahn, and Alsop, 2002). CASL was performed using a 2 sec labeling pulse with amplitude modulated control optimized for 3.0T, with a 0.16 G/cm gradient and 22.5 mG RF irradiation and a low frequency modulation for the control RF of 100Hz (Wang et al., 2005). The post-labeling delay was 1.2 sec.

We used a single shot 3D-GRASE sequence for readout. Imaging parameters were: resolution = $3.9 \times 3.9 \times 6$ mm³, FOV = 250 (H/F) $\times$ 172 (A/P) $\times$ 120 (R/L) mm³, 20 nominal partitions with 10% oversampling, 5/8 partial Fourier encoding, measured partitions = 14, matrix size = 64×44 , BW = 2790 Hz/pixel, gradient-echo spacing = 0.4 msec (with ramp sampling), spin-echo spacing = 23 msec, total read-out time = 340 msec, effective= 70 msec, refocusing flip angle = 180° and TR = 9 sec (Fernandez-Seara et al., 2005).

The original images were realigned to the first image using SPM2 from the Wellcome Department of Imaging Neuroscience (<http://www.fil.ion.ucl.ac.uk/spm/software/spm2>). Images were then converted to axial view using Spmalize (Oakes, 2003), and cerebral blood flow (CBF) maps were subsequently computed from each pair of label/control images using an ASL data processing toolbox, ASLtbx with the simple-subtraction option (Wang, in press). We also monitored head movement during the course of the study. Average head movement was: x-axis: 0.35 mm; y-axis: 0.63 mm; z-axis: 1.7 mm; pitch: 1.5 mm; roll: 0.46 mm; yaw: 0.54 mm.

We used a random effects model to analyze these data with SPM2. Each individual subject’s CBF images were registered to the associated anatomical image, normalized to the MNI standard brain contained in SPM2 and spatially smoothed with a 10-mm FWHM Gaussian kernel. A first-level analysis was used for individual subjects, in which the CBF images comparing Block 5 (narrative story description) to Block 4 and Block 3 (baselines) were performed in each subject using a two-sample paired t-test. These contrasts were fed into a second level analysis to examine a group-wide effect with a one-sample t-test. All of the statistical analyses were performed using the batch scripts contained in ASLtbx. The statistical contrasts were converted to z-scores for each compared voxel. We used a statistical threshold of $p < 0.01$ to generate the images. Subsequently, we used an extent criterion of >20 voxels and only included clusters where the peak voxel generated a z-score > 3.09 . We used these sensitive criteria because the baselines matched our target task so closely. Peak voxels were converted from MNI coordinates to Talairach coordinates (Talairach & Tournoux, 1988) using a non-linear transform (<http://imaging.mrc-cbu.cam.ac.uk/imaging/MniTalairach>).

RESULTS

Behavioral results

Means of coded variables for each block can be found in Table 1. The Wilcoxon rank sum test was used to examine differences between blocks. The P-value is reported whenever it is less than 0.05. The narrative production block contained significantly more narrative elements than the picture description block ($p < .001$). Otherwise, there were no significant differences between the coded variables of total words, syntax, or syllables between these two blocks.

Imaging results

Table 2 summarizes the peak activations in each contrast during which narrative speech was produced. Due to relatively large cluster size, the characterization of extent is important. Coordinates represent activation peaks for each cluster and Brodmann areas refer to the anatomical extent of the activated clusters, which are illustrated in Figure 2. Relative to the combined baseline involving description of a single picture (Block 3) and viewing the picture story while repeatedly saying a nonsense word (Block 4), we found significant activation in bilateral inferior prefrontal regions and in left temporal-parietal cortex during continuous narrative speech (Figure 2, Panel A).

Relative to the description of a single picture, we observed bilateral inferior frontal activation during narrative speech. This is illustrated in Figure 2, Panel B. We also observed marginally significant activation in left dorsolateral prefrontal cortex (Brodmann area (BA) 10; peak coordinates $x = -22$, $y = 56$, $z = 18$; z -score = 2.88) and in left ventrolateral visual association cortex (BA 19; peak coordinates: $x = -38$, $y = -92$, $z = -14$; z -score = 2.83).

The contrast of narrative production minus viewing the picture stimuli ordered as a story (while producing a nonsense word) is illustrated in Figure 2, Panel C. This showed activation in a dorsal portion of left inferior frontal cortex and in supplementary premotor cortex. Activation also was observed in bilateral posterolateral temporal and parietal regions, and in right visual association cortex. Additionally, borderline activation was seen in left ventral inferior frontal cortex (BA 47 and 11; peak coordinates: $x = -22$, $y = 18$, $z = -22$; z -score = 2.84) and left ventrolateral temporal-occipital cortex (BA 19; peak coordinates: $x = -40$, $y = -86$, $z = -22$; z -score = 3.00).

DISCUSSION

Narrative production is a complex process involving at least two major components. Phonological, morphological, and semantic processes are important for the production of single words and the expression of story content. Story content itself is supported by perisylvian language regions and visual and multimodal association cortices. Forming these words into a coherent sentence requires grammatical processing, and combining sentences into a narrative that conveys meaning calls for the recruitment of the higher level processes of planning and organization. These linguistic and cognitive processes are organized in a top-down manner to accommodate the adaptive flexibility needed to produce a coherent narrative (Cooper & Shallice, 2000; Badre & Wagner, 2004). Higher-level organizational processes including working memory are represented, in part, in frontal cortical regions. We designed the current study with the goal of delineating these components, particularly the top-down organizational component of narrative production.

Our observations are consistent with the hypothesis that narrative production depends in part on a network involving frontal and temporal brain regions. Narrative speech production elicits activation in inferior frontal cortex bilaterally, as well as dorsal inferior frontal and lateral temporal-parietal regions of the left hemisphere. Because our baselines were designed to elicit

the lower level processes controlling speech output, we hypothesize that inferior frontal activation seen in our target condition are related specifically to the top-down organization required to produce an extended narrative. Dorsal portions of inferior frontal cortex are likely to support verbal working memory. Bilateral temporal-parietal and visual association cortex activation during narrative production is hypothesized to play a role in the semantic memory required to support inferences derived from long-term memory, that is, extrapolating beyond the meaning available from a stimulus picture. We discuss each of these narrative components below.

Bilateral inferior frontal cortex as an organizational resource

An important component of a narrative is its top-down organization which helps manage an extended story represented in a sequence of pictures. The observation of bilateral inferior frontal activation during narrative production in the present study suggests that this brain region may play a role as a top-down organizational resource rather than support material-specific stimuli such as language (Crozier et al., 1999, Ferstl & von Cramon, 2001; Koechlin, Corrado, Pietrini, & Grafman, 2000; Ramnani & Owen, 2004). The closely matched baselines that we used help us specify the contribution of the activated cortical regions we observed during narrative speech production. In our study, the speech samples elicited during narrative description of an entire story and description of individual unordered pictures thus differed only in terms of the narrative discourse component of the speech sample. This suggests an important role for this inferior frontal region in narrative discourse. Another carefully controlled study of discourse organization by Kuperberg et al (2006) reported frontal activation during coherence judgments rather than narrative production. They attributed frontal recruitment to the controlled activation of semantic information, as described in studies involving single words and pictures (Badre & Wagner, 2004; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997; Thompson-Schill, Aguirre, D'Esposito, & Farah, 1999; Wagner, Pare-Blagoev, Clark, & Poldrack, 2001).

Braun et al (2001) observed left-lateralized activation during their PET study of discourse. The authors attributed their left-lateralized inferior frontal activation to later stages of articulatory encoding and planning for speech production. Borderline activation restricted to left inferior frontal cortex during narrative story expression relative to viewing the story while repeating a nonsense syllable may also reflect planning for speech expression (Braun et al, 2001). However, an account based on articulatory planning would not explain the presence of left inferior frontal activation during narrative expression relative to describing a single picture, where the same level of articulatory planning presumably is required. Regardless of the role of left inferior frontal activation during narrative expression, the absence of right inferior frontal activation during narrative production relative to story viewing emphasizes the role of this region in narrative organization of the visual representation of a story (Ash et al, 2006; Cosentino et al, 2006).

Other work has interpreted inferior frontal activation as voluntary regulation interposed on speech (Schulz, Varga, Jeffries, Ludlow, & Braun, 2005). Our narrative production block and the baseline involving the description of a single picture differ in the higher-level demands important for planning and executing a coherent narrative, but they do not differ in the voluntary regulation of speech. Thus, our observations are more consistent with the hypothesis that this region serves as an organizational resource for narrative production.

Left inferior frontal support for working memory

We observed activation of other areas compared to the combined baseline and the baseline involving story viewing. For example, the dorsal portion of left inferior frontal cortex was recruited. This area was reported by Braun et al (2001) in their PET study of discourse, and by

Kuperberg et al (2006) in their study of discourse judgments. This may be related to the working memory component needed to maintain elements of the narrative in an active state during its production. Evidence consistent with a role for working memory during narrative production comes from superior performance (Singer & Ritchot, 1994) and greater frontal activation (Fiebach, Vos, & Friederici, 2004; Virtue, Haberman, Clancy, Parrish, & Jung Beeman, 2006) in subjects with higher working memory spans. Working memory in the service of complex sentence processing also is associated with dorsal inferior frontal activation (Cooke et al, 2006). Since the story narration block was matched to the picture description block for syllables produced, we do not believe this difference is due to a motor component of speech production. While we cannot rule out the possibility that there are motor differences between the narrative production block and the story viewing block, as supplementary motor area activation is associated with rate differences, we do not find recruitment of other regions also associated with rate effects (Price, Moore, & Frackowiak, 1996). Thus, we believe that dorsal left inferior frontal cortex is recruited as a working memory resource in order to maintain various elements of a narrative in an active state.

Association cortex supports story content and semantic extrapolation

Visual association cortex was activated more prominently during narrative story production relative to either picture description or story viewing. One possibility is that these regions are recruited to decode the visual stimuli depicting the story. Braun et al (2001) and Kuperberg et al (2006) also observed this distribution of temporal-occipital activation in their studies of discourse, even though no picture stimuli were involved, and related this posterior activation to a semantic component of the story. Given the closely matched baselines in the present study, we endorse Kuperberg's (2006) specific proposal that this activation may be related in part to the inferences about meaning that are derived from an extended narrative, that is, beyond the meaning available from direct inspection of a stimulus picture or sentence. This interpretation is also consistent with work showing visual association cortex recruitment during semantic memory challenges (Martin, Ungerleider & Haxby, 2000).

We also observed posterolateral temporal-parietal activation that was more prominent during narrative speech production contrasted with the combined baseline and compared to the baseline of viewing pictures in a story order. We did not find activation of this area relative to the baseline requiring the description of a single picture. This left posterolateral temporal brain region is also associated with processing information represented in semantic memory, regardless of modality (Grossman et al., 2002; Koenig et al., 2005). We speculate that this region works together with visual association cortex to support inferential meaning that is associated with a narrative. Since this is close to auditory association cortex, one possibility is that this activation is related in part to the speaker hearing his or her own speech during narrative production. The activation may have been significant in the left hemisphere because the speech was meaningful. However, this would not explain activation in this area during the comprehension study of Kuperberg et al (2006) where there was no auditory stimulation. We therefore believe the interpretation that temporal-parietal activation indicates a semantic support region important for making story inferences beyond what is available in the picture.

Important Methodological Considerations

Several caveats concerning our imaging acquisition method must be kept in mind during interpretation of our observations. Although CASL perfusion fMRI is an outstanding technique well suited to the collection of brain imaging data over an extended period of time, this technique generally provides lower signal-to-noise ratio for detecting task activation than BOLD fMRI, along with lower spatial and temporal resolution. A new sequence developed after the completion of this study uses background suppression of static brain signal and dramatically improves the sensitivity of ASL perfusion fMRI (Fernandez-Seara et al., 2007).

The high level baselines we used were well suited for this study of narrative organization, but additional work is needed to establish the neural substrate associated with syntactic organization, lexical and phonologic selection, and phonetic planning during narrative production.

Conclusion

With these caveats in mind, a large network appears to be recruited during speech production to help organize and express a coherent narrative. This includes inferior frontal cortex bilaterally, as well as dorsal frontal, temporal-parietal, and temporal-occipital regions of the left hemisphere. Inferior frontal regions appear to be important for supporting the organizational component of a narrative, while dorsal frontal areas may support working memory; temporal-parietal-occipital regions may be associated with the inferential meaning of the narrative.

Acknowledgements

This work was supported in part by NIH (AG17586, AG15116, NS44266, NS53488, RR002305, BCS0517935, and the Dana Foundation).

References

- Aguirre GK, Detre JA, Zarahn E, Alsop DC. Experimental Design and the Relative Sensitivity of BOLD and Perfusion fMRI. *Neuro Image* 2002;15(3):488–500. [PubMed: 11848692]
- Ash S, Moore P, Antani S, McCawley G, Work M, Grossman M. Trying to tell a tale: Discourse impairments in progressive aphasia and frontotemporal dementia. *Neurology* 2006;66:1405–1413. [PubMed: 16682675]
- Badre D, Wagner AD. Selection, integration, and conflict monitoring: Assessing the nature and generality of prefrontal cognitive control mechanisms. *Neuron* 2004;41:473–487. [PubMed: 14766185]
- Belin P, Zilbovicius M, Crozier S, Thivard L, Fontaine A, Masure MC, et al. Lateralization of speech and auditory temporal processing. *Journal of Cognitive Neuroscience* 1998;10(4):536–540. [PubMed: 9712682]
- Birn RM, Cox RW, Bandettini PA. Experimental designs and processing strategies for fMRI studies involving overt verbal responses. *Neuroimage* 2004;23:1046–1058. [PubMed: 15528105]
- Braun AR, Guillemin A, Hosey L, Varga M. The neural organization of discourse: An H215O-PET study of narrative production in English and American sign language. *Brain* 2001;124:2028–2044. [PubMed: 11571220]
- Chao LL, Haxby J, Martin A. Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience* 1999;2:913–919.
- Chapman SB, Bonte FJ, Chiu Wong SB, Zientz J, Hyman LS, Harris TS, et al. Convergence of connected language and SPECT in variants of frontotemporal lobar degeneration. *Alzheimer's Disease and Associated Disorders* 2005;19:202–213.
- Chapman SB, Zientz J, Weiner M, Rosenberg R, Frawley W, Burns MH. Discourse changes in early Alzheimer disease, mild cognitive impairment, and normal aging. *Alzheimer's Disease and Associated Disorders* 2002;16:177–186.
- Cooke A, Grossman M, DeVita C, Gonzalez-Atavales J, Moore P, Chen W, Gee J, Detre J. Large-scale neural network for sentence processing. *Brain and Language* 2006;96:14–36. [PubMed: 16168473]
- Cooper R, Shallice T. Content scheduling and the control of routine activities. *Cognitive Neuropsychology* 2000;17:297–338.
- Cosentino S, Chute D, Libon DJ, Moore P, Grossman M. How Does the Brain Support Script Comprehension? A Study of Executive Processes and Semantic Knowledge in Dementia. *Neuropsychology* 2006;20:307–318. [PubMed: 16719624]
- Crozier S, Sirigu A, Lehericy S, van de Moortele PF, Pillon B, Grafman J, et al. Distinct prefrontal activations in processing sequence at the sentence and script level: An fMRI study. *Neuropsychologia* 1999;37:1469–1476. [PubMed: 10617267]

- Fernández-Seara MA, Wang Z, Wang J, Rao H, Guenther M, Feinberg DA, et al. Continuous arterial spin labeling perfusion measurements using single shot 3D GRASE at 3T. *Magnetic Resonance in Medicine* 2005;54:1241–1247. [PubMed: 16193469]
- Fernández-Seara MA, Wang J, Wang Z, Korczykowski M, Guenther M, Feinberg DA, Detre JA. Imaging mesial temporal lobe activation during scene encoding: Comparison of fMRI using BOLD and arterial spin labeling. *Human Brain Mapping*. 2007;in press
- Ferstl EC, Guthke T, von Cramon DY. Change of perspective in discourse comprehension: encoding and retrieval processes after brain injury. *Brain and Language* 1999;70:385–420. [PubMed: 10600227]
- Ferstl EC, von Cramon DY. The role of coherence and cohesion in text comprehension: An event-related fMRI study. *Cognitive Brain Research* 2001;11:325–340. [PubMed: 11339984]
- Ferstl EC, von Cramon DY. What does the frontomedial cortex contribute to language processing: Coherence or theory of mind? *Neuroimage* 2002;17:1599–1612. [PubMed: 12414298]
- Fiebach CJ, Vos SH, Friederici AD. Neural correlates of syntactic ambiguity in sentence comprehension for low and high span readers. *Journal of Cognitive Neuroscience* 2004;16:1562–1575. [PubMed: 15601519]
- Fletcher PC, Happe F, Frith U, Baker SC, Dolan RJ, Frackowiak RSJ, et al. Other minds in the brain: a functional imaging study of “theory of mind” in story comprehension. *Cognition* 1995;57:109–128. [PubMed: 8556839]
- Gallagher H. Reading the mind in cartoons and stories: An fMRI study of ‘theory of mind’ in verbal and nonverbal tasks. *Neuropsychologia* 2000;38:11–21. [PubMed: 10617288]
- Glosser G, Deser T. Aging changes in microlinguistic and macrolinguistic aspects of discourse production. *Journal of Gerontology: Psychological Sciences* 1992;47:266–272.
- Grossman M, McMillan C, Moore P, Ding L, Glosser G, Work M, et al. What’s in a name: Voxel-based morphometric analyses of MRI and naming difficulty in Alzheimer’s disease, frontotemporal dementia, and corticobasal degeneration. *Brain* 2004;127:628–649. [PubMed: 14761903]
- Grossman M, Smith EE, Koenig P, Glosser G, DeVita C, Moore P, et al. The neural basis for categorization in semantic memory. *Neuroimage* 2002;17:1549–1561. [PubMed: 12414293]
- Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR, Gurney EM, Bowtell RW. “Sparse” temporal sampling in auditory fMRI. *Hum Brain Mapp* 1999;7:213–223. [PubMed: 10194620]
- Kemeny S, Xu J, Park GH, Hosey LA, Wettig CM, Braun AR. Temporal dissociation of early lexical access and articulation using a delayed naming task -- An fMRI study. *Cerebral Cortex* 2006;16:587–595. [PubMed: 16049190]
- Kemeny S, Ye FQ, Birn RM, Braun AR. Comparison of continuous overt speech fMRI using BOLD and arterial spin labeling. *Human Brain Mapping* 2005;24:173–183. [PubMed: 15486986]
- Kim J, Whyte J, Wang J, Rao H, Tang KZ, Detre JA. Continuous ASL perfusion fMRI investigation of higher cognition: Quantification of tonic CBF changes during sustained attention and working memory tasks. *Neuroimage* 2006;31:376–385. [PubMed: 16427324]
- Koechlin E, Corrado G, Pietrini P, Grafman J. Dissociating the role of the medial and lateral anterior prefrontal cortex in human planning. *Proceedings of the National Academy of Sciences of the USA* 2000;97:7651–7656. [PubMed: 10852964]
- Koenig P, Smith EE, Glosser G, DeVita C, Moore P, McMillan C, et al. The neural basis for novel semantic categorization. *Neuroimage* 2005;24:369–383. [PubMed: 15627580]
- Kuperberg G, Lakshmanan BM, Caplan D, Holcomb PJ. Making sense of discourse: An fMRI study of causal inferencing across sentences. *Neuroimage* 2006;33:343–361. [PubMed: 16876436]
- Libon DJ, Massimo L, Moore P, Coslett HB, Chatterjee A, Aguirre GK, et al. Differentiating the frontotemporal dementias from Alzheimer’s disease: The Philadelphia Brief Assessment of Cognition. 2007;submitted
- Mar RA. The neuropsychology of narrative: Story comprehension, story production, and their interrelation. *Neuropsychologia* 2004;42:1414–1434. [PubMed: 15193948]
- Martin, A.; Ungerleider, L.; Haxby, JV. Category specificity and the brain: The sensory/motor model of semantic representations of objects. In: Gazzaniga, MS., editor. *The New Cognitive Sciences*. Cambridge: MIT Press; 2000. p. 1023–1036.

- Mason RA, Just MA. How the brain processes causal inferences in text. *Psychological Science* 2004;15:1–7. [PubMed: 14717824]
- Mayer, M. *Frog, Where Are You?*. New York: Penguin Books; 1969.
- Mazoyer BM, Tzourio N, Frak V, Syrota A, Murayama N, Levrier O, et al. The cortical representation of speech. *Journal of Cognitive Neuroscience* 1993;5:467–479.
- Munhall K. Functional imaging during speech production. *Acta Psychologica* 2001;107:95–117. [PubMed: 11388145]
- Oakes, T. SPAMALIZE: canned data analysis software. 2003. http://brainimaging.waisman.wisc.edu/oakes/spam/spam_frames.htm
- Palmer E, Rosen H, Ojemann J, Buckner RL, Kelley WM, Petersen S. An event-related fMRI study of overt and covert word stem completion. *Neuroimage* 2001;14:182–193. [PubMed: 11525327]
- Price CJ, Moore CJ, Frackowiak RSJ. The Effect of Varying Stimulus Rate and Duration on Brain Activity during Reading. *NeuroImage* 1996;3:40–52. [PubMed: 9345474]
- Ramnani N, Owen AM. Anterior prefrontal cortex: Insights into function from anatomy and neuroimaging. *Nature Reviews: Neuroscience* 2004;5:184–194.
- Robertson DA, Gernsbacher MA, Guidotti SJ, Robertson RR, Irwin W, Mock BJ, et al. Functional neuroanatomy of the cognitive process of mapping during discourse comprehension. *Psychological Science* 2000;11:255–260. [PubMed: 11273413]
- Rosen H, Ojemann J, Ollinger J, Petersen S. Comparison of word retrieval done silently and aloud using fMRI. *Brain and Cognition* 2000;42:201–217. [PubMed: 10744920]
- Rosen HJ, Allison SC, Schauer GF, Gorno-Tempini ML, Weiner MW, Miller BL. Neuroanatomical correlates of behavioural disorders in dementia. *Brain* 2005;128:2612–2625. [PubMed: 16195246]
- Schulz GM, Varga M, Jeffries K, Ludlow CL, Braun AR. Functional Neuroanatomy of Human Vocalization: An H215O PET Study. *Cerebral Cortex* 2005;15:1835–1847. [PubMed: 15746003]
- Shuster L, Lemieux S. An fMRI investigation of covertly and overtly produced mono- and multisyllabic words. *Brain and Language* 2005;93:20–31. [PubMed: 15766765]
- Singer M, Ritchot KF. The role of working memory capacity and knowledge access in text inference processing. *Memory and Cognition* 1994;24:733–743.
- Soros P, Sokoloff LG, Bose A, McIntosh AR, Graham SJ, Stuss D. Clustered functional MRI of overt speech production. *Neuroimage* 2006;32:376–387. [PubMed: 16631384]
- Talairach, J.; Tournoux, P. *Co-planar stereotaxic atlas of the human brain*. 1. New York: Thieme Medical Publishing Company; 1988.
- Thompson-Schill SL, Aguirre G, D'Esposito M, Farah MJ. A neural basis for category and modality specificity of semantic knowledge. *Neuropsychologia* 1999;37:671–676. [PubMed: 10390028]
- Thompson-Schill SL, D'Esposito M, Aguirre G, Farah MJ. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences of the USA* 1997;94:14792–14797. [PubMed: 9405692]
- Virtue S, Haberman J, Clancy Z, Parrish TB, Jung Beeman M. Neural activity of inferences during story comprehension. *Brain Research* 2006;1084:104–114. [PubMed: 16574079]
- Wagner AD, Pare-Blagoev EJ, Clark J, Poldrack RA. Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. *Neuron* 2001;31:329–336. [PubMed: 11502262]
- Wang J, Aguirre GK, Kimberg DY, Roc AC, Li L, Detre JA. Arterial spin labeling perfusion fMRI with very low task frequency. *Magnetic Resonance in Medicine* 2003;49:796–802. [PubMed: 12704760]
- Wang J, Zhang Y, Wolf RL, Roc AC, Alsop D, Detre J. Amplitude-modulated continuous arterial spin-labeling 3.0-T perfusion MR imaging with a single coil: Feasibility study. *Radiology* 2005;235:218–228. [PubMed: 15716390]
- Wang Z, Aguirre GK, Rao H, Wang J, Fernández-Seara MA, Childress AR, Detre JA. Empirical optimization of ASL data analysis using the ASL data processing toolbox: ASLtbx: ALSLtbx. *Magnetic Resonance Imaging*. in press
- Wildgruber D, Ackermann H, Grodd W. Differential contributions of motor cortex, basal ganglia, and cerebellum to speech motor control: effects of syllable repetition rate evaluated by fMRI. *Neuroimage* 2001;13:101–109. [PubMed: 11133313]

- Wildgruber D, Ackermann H, Klose U, Kardatzki B, Grodd W. Functional lateralization of speech production at primary motor cortex: a fMRI study. *NeuroReport* 1996;7:2791–2795. [PubMed: 8981469]
- Williams GB, Nestor PJ, Hodges JR. Neural correlates of semantic and behavioural deficits in frontotemporal dementia. *Neuroimage* 2005;24:1042–1051. [PubMed: 15670681]
- Xu J, Kemeny S, Park G, Frattali C, Braun A. Language in context: emergent features of word, sentence, and narrative comprehension. *Neuroimage* 2005;25:1002–1015. [PubMed: 15809000]

FIGURE 1. TIMELINE OF STIMULUS PRESENTATION

FIGURE 2. SIGNIFICANT ACTIVATIONS DURING NARRATIVE SPEECH PRODUCTION
 PANEL A: NARRATIVE SPEECH PRODUCTION COMPARED TO A COMBINED
 BASELINE OF SINGLE PICTURE DESCRIPTION AND STORY VIEWING
 PANEL B: NARRATIVE SPEECH PRODUCTION COMPARED TO SINGLE PICTURE
 DESCRIPTION
 PANEL C: NARRATIVE SPEECH PRODUCTION COMPARED TO STORY VIEWING

Table 1

Summary of Coded Variables for Speech Samples

	Syllables	Words	Syntax	Narrative Element
Picture Description Block	15.3±2.9	13.1±2.3	2.1±0.3	0.04±0.05
Narrative Block	16.5±2.2	14.1±1.7	2.2±0.2	0.51±0.1

Table 2

Activations during Narrative Speech

Anatomic area	Brodmann area	Talairach coordinates			Z-score
		x	y	z	
Narrative Production minus Single Picture Description & Story Viewing (During Nonsense Word Production)					
Left temporal-parietal	39, 40	-32	-44	54	3.35
Left dorsal inferior frontal	6	-38	4	46	3.27
Bilateral ventral inferior frontal	47, 11	22	38	-24	3.06
Narrative Production minus Single Picture Description					
Bilateral ventral inferior frontal	47, 11	14	20	-28	3.40
Narrative Production minus Story Viewing (During Nonsense Word Production)					
Left dorsal inferior frontal	44, 6	-40	2	48	4.09
Left temporal-parietal	39, 40	-26	-74	18	3.73
Left middle temporal	21, 22	-34	-72	0	3.73
Right temporal-parietal	39, 22	34	-78	26	3.46
Right inferior temporal-occipital	19	38	-70	-14	3.15