Question/Statement Judgments: An fMRI Study of Intonation Processing

Colin P. Doherty,[1]* W. Caroline West,[1] Laura C. Dilley,[2] Stefanie Shattuck-Hufnagel,[2] and David Caplan[1]

[1]*Neuropsychology Laboratory and MGH/MIT/HMS Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Boston, Massachusetts*
[2]*Speech Communication Group, Massachusetts Institute of Technology, Cambridge, Massachusetts*

Abstract: We examined changes in fMRI BOLD signal associated with question/statement judgments in an event-related paradigm to investigate the neural basis of processing one aspect of intonation. Subjects made judgments about digitized recordings of three types of utterances: questions with rising intonation (RQ; e.g., "She was talking to her father?"), statements with a falling intonation (FS; e.g., "She was talking to her father."), and questions with a falling intonation and a word order change (FQ; e.g., "Was she talking to her father?"). Functional echo planar imaging (EPI) scans were collected from 11 normal subjects. There was increased BOLD activity in bilateral inferior frontal and temporal regions for RQ over either FQ or FS stimuli. The study provides data relevant to the location of regions responsive to intonationally marked illocutionary differences between questions and statements. *Hum Brain Mapp 23: 85–98, 2004.* © 2004 Wiley-Liss, Inc.

Key words: intonation; fMRI; question; linguistic prosody

INTRODUCTION

Intonation is one of the three basic elements of sentence prosody, the others being metrical rhythm and prosodic phrasing [Selkirk, 1995]. Intonation refers to the use of suprasegmental phonetic features to convey "post-lexical" or sentence-level meanings and intentionally excludes features of lexical stress, accent, and tone, which serve to distinguish one word from another. Intonation may be used to convey both non-categorical "paralinguistic" contrasts such as emotional states and categorical linguistic contrasts [Ladd, 1996]. One categorical linguistic value that can be conveyed by intonation is the illocutionary force of an utterance. Austin [1975] introduced the notion that human speech can be conceived as consisting of numerous "speech acts." Even when they contain the same words and convey the same relations between their words (i.e., when they have the same propositional meaning), statements and questions differ as speech acts, since a question involves a certain type of intention [Searle, 1969]. The illocutionary act of "posing a question" may be signaled lexically by the use of wh- words, by a change in word order, or prosodically, through the use of intonation [Couper-Kuhlen, 1986].

Despite a substantial number of behavioral and lesion studies reported over the past 40 years, no clear consensus on the neuroanatomical representation of prosody has emerged, other than the categorization of paralinguistic "emotional" prosody as being represented in the right hemisphere [Blumstein and Cooper, 1974; Heilman et al., 1975;

Colin P. Doherty and W. Caroline West contributed equally to this work.

Contract grant sponsor: Grass Foundation, Braintree, MA; Contract grant sponsor: NIDCD; Contract grant number: DC02146; Contract grant sponsor: The National Center for Research Resources; Contract grant number: P41RR14075; Contract grant sponsor: Mental Illness and Neuroscience Discovery (MIND) Institute.

*Correspondence to: Dr. Colin P. Doherty, Senior Lecturer Neuroscience, Royal College of Surgeons in Ireland, Beaumont Hospital, Dublin 9, Ireland. E-mail: cpdoherty@partners.org

Received for publication 2 December 2002; Accepted 25 February 2004
DOI: 10.1002/hbm.20042
Published online in Wiley InterScience (www.interscience.wiley.com).

Ross, 1981; Ross and Mesulam, 1979; Ross et al., 1997; Tucker et al., 1977; Wolfe and Ross, 1987]. Even this conclusion is unsettled; there have been a number of studies that have failed to find a clear distinction between the hemispheres' processing of emotional information [Pell and Baum, 1997; Schlanger et al., 1976; Van Lancker and Sidtis, 1992; Zurif and Mendelsohn, 1972]. Lesion studies looking specifically for the representation of "linguistic prosody" are sparse, but a few have used question-statement judgments as an example of linguistic prosody and have reported deficits due to damage in either hemisphere [Baum et al., 1982; Blumstein and Goodglass, 1972; Behrens, 1988; Shapiro and Danly, 1985]. This lack of clear lateralization has led some authors to suggest a co-operative role for each hemisphere in linguistic prosodic processing [Goodglass and Calderon, 1977; Heilman et al., 1984].

Functional imaging studies have also generally been designed to examine emotional contrasts in speech prosody. The findings of both PET and fMRI studies concur with the majority of lesion data showing that emotionally intoned speech causes preferential activation in the right hemisphere [Buchanan et al., 2000; George et al., 1996; Imaizumi et al., 1997; Mitchell et al., 2003; Wildgruber et al., 2002]. Only a small number of functional studies in normal subjects have specifically addressed the localization of linguistic prosody. Stiller and colleagues found that when monitoring either strings of nonsense syllables or real adjectives, BOLD activity increased bilaterally equally for both phoneme detection and question–statement intonation judgments. This contrasted with BOLD activity being lateralized to the right when the subjects attended to emotional contrasts [Stiller et al., 1997]. This study was limited to a region of interest analysis of the acoustically responsive areas on the supratemporal plane. In a study comparing sentences that were filtered leaving only prosodic information with sentences with grammatically congruent non-words, prosodic processing preferentially activated the right superior temporal region while the grammatical constructions activated the homologous areas in the left hemisphere [Meyer et al., 2002]. Gandour et al. [2003] examined BOLD signal responses to discrimination of differences in illocutionary force (questions vs. statements) and emotional valence (happy vs. angry vs. sad) in Chinese utterances in Chinese and English speakers. In both groups of subjects, discrimination of illocutionary force compared to a passive listening baseline led to widespread increased BOLD signal in frontal, parietal, and temporal lobes bilaterally and comparison of making judgments regarding intonation vs. emotional valence led to bilateral frontopolar and insular activation. Comparison of making judgments regarding intonation vs. emotional valence led to left posterior prefrontal, inferior parietal, and occipitotemporal activation in the Chinese subjects and to right prefrontal activation in the English subjects. Overall, these studies suggest bihemispheric processing of illocutionary intonation.

In this study, we examined changes in fMRI BOLD signal associated with intonation contours that signaled questions and statements. The results of studies addressing the neural basis of linguistic prosody led us to hypothesize that the vascular response to manipulation of phonologically relevant intonation would reflect the activation of a network of cortical areas encompassing primary and secondary association areas on both supratemporal planes. Because of the small number of these studies, we also considered the hypothesis that other areas of the brain that have been implicated in processing speech sounds, phonological representations, phonological short-term memory, and sentence-level semantic representations would also be active in the contrasts we studied. These areas include bilateral inferior frontal, inferior temporal, and inferior parietal areas [for relevant evidence see, among many other references, Belin et al., 2000; Binder et al., 1994, 1996, 1997; Burton et al., 2000; Damasio et al., 1996; Demonet et al., 1992; Fiez et al., 1995; Frith et al., 1991; Gabrieli et al., 1998; Howard et al., 1992; Hickok and Poeppel, 2000; Johnsrude et al., 2000; Kapur et al., 1994; Pardo et al., 1999; Patel and Balaban, 2001; Paulesu et al., 1993; Perani et al., 1996; Poldrack et al., 1999; Price et al., 1996; Vandenberghe et al., 1996; Warburton et al., 1996; Wise et al., 1991; Zatorre et al., 1992, 1994; Zatorre and Belin, 2001]. We were particularly interested in whether there were changes in vascular responses in these areas that are associated with particular intonational contours.

SUBJECTS AND METHODS

Subjects

Eleven normal subjects (four men, seven women; mean age: 23.1 years, range: 18–26; mean years of education: 15.5 years) were recruited for the fMRI study. All were right-handed (based on the Edinburgh handedness inventory [Oldfield, 1971]), native speakers of American English, college educated, and had no history of neurological or psychiatric disease. All gave informed consent and were paid for their participation.

Stimulus Materials

Stimuli consisted of digitized auditory recordings of a human male voice enunciating sentences in standard North American English. The speaker was the North American youth oratory champion in 1999. One hundred and fifty triads of sentences (450 sentences in total), based on concept stems, were constructed. Each sentence triad consisted of a string of lexical items, which were intoned as (1) a question with a rising boundary tone (RQ: "She was talking to her father?"), (2) a statement with a falling boundary tone (FS: "She was talking to her father") or (3) as a question with a word-order change to denote the illocutionary intent resulting in a question with a falling intonation contour (FQ: "Was she talking to her father?"). All sentences consisted of the feminine pronoun and the past progressive tense as the stem (for RQ and FS: "She was") with the order inverted for the FQ sentences ("Was she"). Each sentence was between 1.8 and 2.2 s long. Figure 1, created using the PRAAT software system (PRAAT: doing phonetics by computer, v 4113, 2003, University of Amersterdam, The Netherlands), illustrates the intonation contours created by the three corresponding stimuli of one concept stem.

Three lists of 150 related sentences were constructed in order to counterbalance the three sentence types across subjects. Thus, each subject heard 50 RQ, 50 FS, and 50 FQ sentences with no concept being repeated within subjects. The three sentence types were pseudo-randomized along with white noise segments of varying duration (which served as a fixation condition) to optimize the efficiency of the deconvolution and estimation of the hemodynamic response [Burock et al., 1998; Dale, 1999].

Psycholingusitic Methods

Each sentence was presented during a 4-s epoch with the end of the sentence falling on the 4-s mark such that the final boundary tones were aligned. At the end of the sentence, there was a 500-ms silence followed by a 1,000-ms period during which time the subject was prompted to make a response by the visual command "Respond Now." This was followed by a further 500-ms silence before the next stimulus. Each trial thus lasted 6,000 ms. The subjects' task was to indicate whether each sentence was a question or a statement.

Stimuli (sentences and white noise) were presented in 5 blocks. Each block lasted 240 s. The total experimental length (including 2 min rest after each block) was 30 min. The stimuli were presented using the E-prime v.1.0 software package (Psychology Software Tools, Inc., Pittsburgh, PA), a PC based presentation program that was loaded on a 850-MHz Pentium III Dell laptop computer.

Emotional intensity study

Because intonation can convey emotional expression, which might be confounded with the illocutionary values we manipulated in this study, a behavioral study was performed to evaluate the emotional intensity of the stimuli. Eight subjects were presented the same stimuli as the imaging group using standard stereo headphones while watching a fixation module on the screen of the laptop. Subjects were asked to rate the emotionality of each stimulus on a 7-point scale ranging from highly unemotional (1) to highly emotional (7). We explained that "emotionality" referred to emotional states such as sadness, anger, and happiness. The subjects scored each stimulus by pressing the appropriate score on the numeric keyboard during the interstimulus interval.

fMRI study

Each subject had a pair of electrostactic headphones placed comfortably over both ears. Surrounding this was a tight-fitting helmet of reinforced neoprene rubber designed to reduce extraneous scanner noise. The subject's head was then immobilized with foam pillows to reduce motion artifact. During stimulus presention (sentences and white noise), subjects viewed a white fixation point on a black screen projected at the rear of the magnet bore using a shielded LCD projector (Sharp 2000 Color LCD projector) through a mirror attached to the headcoil. The crosshairs were replaced with the words "Respond Now" at the end of the spoken stimuli. For the white noise stimuli, which varied in length according to the optimization procedure, the crosshairs remained during the entire presentation.

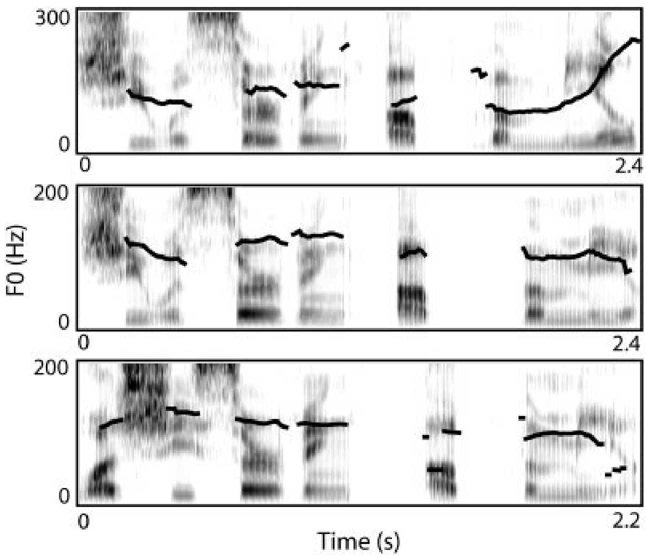

Figure 1.
Three paired stimuli illustrating the different intonation conditions in the experiment. The data set consists of voice spectrographs with uncorrected fundamental frequency (pitch) contours superimposed as a dark black line. From top to bottom: an RQ (Rising Question) utterance (*She was serving up the meal?*), an FS (Falling Statement) utterance (*She was serving up the meal*), and an FQ (Falling Question) utterance (*Was she serving up the meal*). (Figure created using PRAAT software.)

Subjects made question/statement judgements by pressing a key on a custom-designed magnet-compatible button box placed at the level of the hip to record the subject's performance. Subjects used their left hand to signal questions and their right hand to signal statements. At the start of the experiment, one 88-s block consisting of 12 practice sentences and 16 s of noise was used to assess the ability to hear and discriminate the sentences.

Magnetic Resonance Imaging Procedure

Subjects were scanned in two separate sessions. In both scanning sessions, subjects were placed in the standard Siemens quadrature head coil. In the structural session, two sets of high-resolution anatomical images were acquired in a 1.5T whole-body Siemens Sonata scanner (Siemens Medical Systems, Iselin, NJ) using a T1-weighted MP-RAGE sequence (TR = 7.25 ms, TE = 3.0 ms, and flip angle = 7°). Volumes consisted of 128 sagittal slices (effective thickness = 1.33 mm; matrix size = 192 × 256; FOV = 256 mm; in-plane resolution = 1.33 × 1.0 mm).

The functional session utilized a 3.0T head-only Siemens Allegra scanner (Siemens Medical Systems). Scans included a lower-resolution T1-weighted MP-RAGE sequence, (TR = 11.08 ms, TE = 4.3 ms, flip angle = 8°) which consisted of 128 sagittal slices (effective slice thickness = 1.33 mm, matrix size = 128 × 256, FOV = 256 mm, in-plane resolution = 2.0 × 1.0 mm). A T1-weighted inversion-recovery echo-planar sequence (TR = 6 s, TE = 29 ms, flip angle = 90°) with 33 slices aligned parallel to the line defined by the anterior- and posterior-commissures (AC-PC) was also acquired to aid registration of the functional images with the high-resolution anatomical images. The slices were effectively 3 mm thick and had a distance of 0.9 mm between slices. The in-plane resolution was 3.13 × 3.13 mm (64 × 64 matrix, 200 mm FOV).

The functional volume acquisitions utilized a T2*-weighted gradient-echo pulse sequence (TR = 2 s, TE = 25 ms, and flip angle = 90°). The volume was comprised of 33 transverse slices aligned along the same AC-PC plane as the registration volume. The interleaved slices were effectively 3 mm thick with a distance of 0.9 mm between slices. The in-plane resolution was 3.13 × 3.13 mm (64 × 64 matrix, 200 mm FOV). Each run consisted of 120 such volume acquisitions for a total of 3,960 images. The 33 slices of a single volume took the entire TR (2 s) to be fully acquired and a new volume was initiated every TR by definition. An initial 8-s (4 TR equivalent) buffer of RF pulse activations, during which no stimulus items were presented and no functional volumes were acquired, was employed to ensure maximal signal during the length of the functional run.

Cortical Surface Reconstruction

The high-resolution anatomical MP-RAGE scans were used to construct a model of each subject's cortical surface. An average of the two structural scans was used to maximize the signal to noise ratio. The cortical reconstruction procedure involved (1) segmentation of the cortical white matter; (2) tessellation of the estimated border between gray and white matter, providing a geometrical representation for the cortical surface of each subject; and (3) inflation of the folded surface tessellation to unfold cortical sulci, allowing visualization of cortical activation in both the gyri and sulci simultaneously [Dale et al., 1999; Fischl et al., 1999a, 2001].

For purposes of inter-subject averaging, the reconstructed surface for each subject was morphed onto an average spherical representation. This procedure optimally aligns sulcal and gyral features across subjects, while minimizing metric distortion, and establishes a spherical-based co-ordinate system onto which the selective averages and variances of each subject's functional data can be resampled [Fischl et al., 1999a,b]. This non-rigid, surface-based deformation procedure results in a substantial reduction in anatomical and functional variability across subjects relative to the more commonly used normalization approach of Talairach [Talairach and Tournoux, 1988], thereby improving the anatomical precision of the inter-subject averages.

Functional Pre-processing

Pre-processing and statistical analysis of the functional MRI data were performed using the FreeSurfer Functional Analysis Stream (FS-FAST) developed at the Martinos Center [Burock and Dale, 2000]. For each subject, the acquired native functional volumes were first corrected for potential motion of the participant using the AFNI algorithm [Cox, 1996]. Next, the functional volumes were spatially smoothed using a three-dimensional technique and a full-width half-max (FWHM) of 6 mm. Global intensity variations across runs and subjects were removed by rescaling all voxels and time points of each run such that the mean in-brain intensity was fixed at an arbitrary value of 1,000.

The functional images for each subject were analyzed with a General Linear Model (GLM) using a finite impulse response model (FIR) of the event-related hemodynamic response [Burock and Dale, 2000]. The FIR gives an estimate of the hemodynamic response average at each TR within a peri-stimulus window. The FIR does not make any assumption about the shape of the hemodynamic response. Mean offset and linear trend regressors were included to remove low-frequency drift. The autocorrelation function of the residual error, averaged across all brain voxels, was used to compute a global whitening filter in order to account for the intrinsic serial autocorrelation in fMRI noise. The GLM parameter estimates and residual error variances of each subject's functional data were resampled onto his or her inflated cortical surface and into the spherical coordinate system using the surface transforms described above. Each subject's data were then smoothed on the surface tessellation using an iterative nearest-neighbor averaging procedure equivalent to applying a two-dimensional Gaussian smoothing kernel with a FWHM of approximately 8.5 mm. Because this smoothing procedure was restricted to the cortical surface, averaging data across sulci or outside gray matter was avoided.

Statistical Analyses

Voxelwise analysis

Contrasts of interest were tested at each voxel on the spherical surface across the group using a random effects model of the cross-subject variance of the FIR parameter estimates. Contrasts were constructed over a range of post-stimulus delays in the FIR model corresponding to the delays at which vascular responses were expected to be associated with intonational and illocutionary perception. BOLD signal changes follow electrophysiological events associated with elementary sensory stimuli and simple motor functions by as little as 2 seconds, with an established response by 4–6 s [Bandettini, 1993; Turner et al., 1997]. Thus, with the end of the auditory sentence occurring at 4 s, the vascular response to the auditory perception of sentence-final intonational change would be expected to start by 6 s and to be established by 10 s. Thus, the BOLD signal was collapsed across the three post-stimulus delay intervals of 6, 8, and 10 s.

Group statistical activation maps were constructed for contrasts of interest using a *t* statistic. To correct for multiple comparisons, we identified significant clusters of activated voxels on the basis of a Monte Carlo simulation, as follows. A volume of Gaussian distributed numbers was generated for each subject, and was processed in the same manner as the real data, including volumetric smoothing, resampling onto the sphere, smoothing on the spherical surface, random effects analysis, and significance map generation. A clustering program was run on these maps to extract clusters of voxels whose members each exceeded a specified threshold and whose area was equal to or greater than a specified size. This process was repeated 3,500 times, allowing us to compute the likelihood of getting one or more clusters of a given size and voxel threshold under the null hypothesis. We set the threshold for cluster size at 300 mm^2 and the threshold for rejection of the null hypothesis at $P < 0.05$. The real data were then subjected to the same clustering as applied to the simulated data. The resulting statistical activation maps are shown in Figure 3. The functional activations are displayed on a map of the average folding patterns of the cortical surface, derived using the surface-based morphing procedure [Fischl et al, 1999a,b].

Region of interest analysis

In a region of interest (ROI)-based approach to data analysis, 17 hypothesis-driven anatomical ROIs in each hemisphere, corresponding to frontal, parietal, and temporal regions considered to be involved in acoustic/phonetic, phonological, and semantic processing were defined on the average cortical surface in accordance with the MGH Center for Morphometric Analysis (CMA) parcellation system [Caviness et al, 1996; Rademacher et al., 1992]. These ROIs are shown in Figure 2. These ROIs were then translated onto each individual subject's surface using the transformation matrices generated during the morphing procedure described above. For each subject, the mean percent signal change within each ROI relative to the prestimulus baseline was calculated for each experimental condition at each TR. As in the voxelwise analysis, the values at post-stimulus delays of 6, 8, and 10 s were averaged to yield a single percent signal change value for each condition. The resulting data were analyzed in SAS (SAS/STAT software V6) in an ANOVA for repeated measures with factors of hemisphere (2), ROI (17), and sentence type (3). Further analysis of significant main effects and interactions was performed using Tukey's least mean squares adjustment for multiple comparisons with a significance level set at $P < 0.05$.

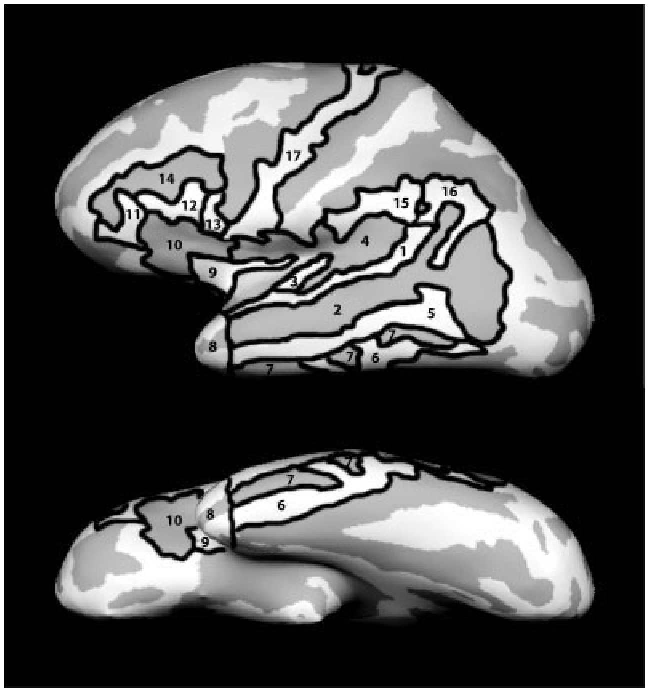

Figure 2.
Schematic map of the 17 regions of interest defined on the averaged folding patterns of the cortical surface. The lateral surface of the left hemisphere is depicted in the top view and the ventral surface is depicted in the bottom view. Light gray areas represent gyri. Dark gray areas represent sulci. (1) superior temporal gyrus, (2) superior temporal sulcus, (3) Heschl's gyrus, (4) Sylvian fissure, (5) middle temporal gyrus, (6) inferior temporal gyrus, (7) inferior temporal sulcus, (8) temporal pole, (9) insula, (10) circular sulcus of insula, (11) inferior frontal gyrus-pars orbitalis, (12) inferior frontal gyrus-pars triangularis, (13) inferior frontal gyrus-pars opercularis, (14) inferior frontal sulcus, (15) supramarginal gyrus, (16) angular gyrus, (17) precentral gyrus.

RESULTS

Behavioral Results

Emotional intensity ratings

The mean and standard deviations for the emotionality ratings (7-point scale) for RQ, FQ, and FS were 4.1 ± 0.4, 3.9 ± 0.1, and 4 ± 0.1 respectively. An analysis of variance (ANOVA) revealed no significant main effects of sentence type (F (2, 14) = 1.7, $P = 0.24$).

FMRI study: identification accuracy

Behavioral data from one subject were not recorded. The mean percentage correct and SD from the remaining 10 subjects for identification accuracy of RQ, FQ, and FS were 92.8 ± 1.9%, 91.4 ± 1.9%, and 94.2 ± 2.1% respectively. An ANOVA revealed no significant main effect of sentence type (F (2, 18) = 1.4, $P = 0.3$).

Imaging Results

Overall cortical activation (voxelwise analysis)

Figure 3 shows the statistical activation maps of the group averaged data superimposed upon a map of the averaged folding patterns of the cortical surface. Table I shows the significance level of the peak activation in each region and the corresponding Talairach coordinates as well as the size (mm^2) of each cluster.

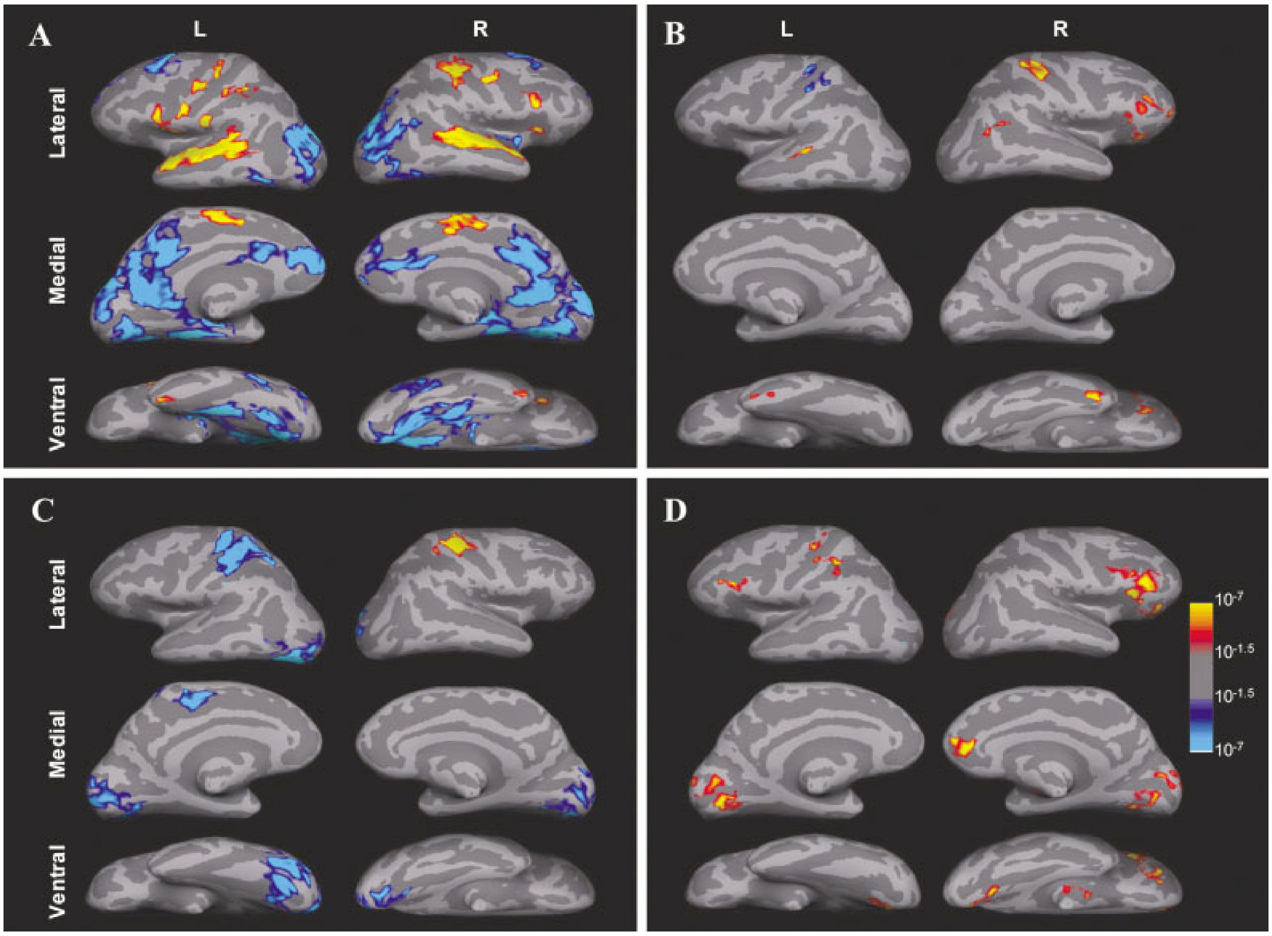

Figure 3.

Statistical activation maps of the group averaged data superimposed upon a map of the averaged folding patterns of the cortical surface for the following contrasts: (**A**) all conditions versus white noise, (**B**) RQ versus FS, (**C**) FQ versus FS, and (**D**) RQ versus FQ. The corrected threshold for significance (the likelihood of getting one or more clusters of at least 300 mm^2 under the null hypothesis) equals $P < 0.05$. The scale bar indicates the P value (log 10) at each voxel.

In Figure 3A, the activity generated by all the sentences regardless of type is compared to the white noise fixation condition. Increased BOLD signal for spoken sentences over white noise (red/yellow) was seen in both superior temporal regions and to a lesser extent in both lateral inferior frontal regions (more anteriorly in pars triangularis on the right and more posteriorly in pars opercularis on the left). There was also increased activation in motor and premotor areas consistent with button push activity. Finally, there was a widespread increase in activation favoring white noise (blue areas) in the lateral and medial occipital areas, medial frontal areas, and inferior temporo-occipital areas. The activity in the medial and polar visual cortical areas regions presumably reflects the visual parts of the task (perhaps the longer period of presentation of the fixation symbol in the white noise condition than the visual stimuli in the experimental conditions) and the activity in more medial and frontal areas presumably reflects a range of factors associated with anticipation and concentration. Based on its location, we do not think this activity is directly related to processing intonation, though it may partially reflect how such processing elsewhere affects a complex neural net. Because we do not have a basis upon which to explain this result, it will not be discussed further.

Figure 3B–D shows the statistical maps generated by paired comparisons of the BOLD activity in each of the main conditions. Increased and corresponding decreased activity seen in motor areas in the precentral gyrus and adjacent parietal sensory cortex in the comparisons across conditions is consistent with left-hand motion in response to questions and right-hand motion in response to statements. FQ utterances produced less occipital activity bilaterally than either RQ or FS utterances; the reasons for this also remain unexplained. The descriptions of activation below do not mention these areas of activation further.

When the BOLD activity generated by Rising Questions (RQ) was compared to Falling Statements (FS) (Fig. 3B), statistically increased activity for RQ over FS was noted in

TABLE I. Areas of activation in the voxelwise analysis

Contrast	Hemi	Surface region	BA	Peak p (log 10)	Size (mm^2)	Talairach coordinates		
						x	*y*	*z*
RQ-FS	LH	Superior temporal gyrus	22	3.3	336	−61	−27	1
		Inferior temporal gyrus	20	2.4	333	−39	−4	−35
RQ-FS	RH	SMG/AG	39/40	2.6	371	48	−53	28
		Inferior temporal gyrus	20	3.6	516	36	−6	−29
		antIFG/frontal pole	10	2.8	1,319	25	58	−3
		IFGorb/orbital gyrus	47/11	2.8	384	33	30	−12
RQ-FQ	LH	Inferior frontal gyrus	45	3.4	555	−52	27	5
RQ-FQ	RH	Anterior cingulate sulcus	32/10	4.5	715	8	40	−1
		Inferior frontal gyrus	45/47	4.0	2,535	38	33	8
		IFGorb/orbital gyrus	47/11	3.1	669	27	36	−5

the left hemisphere in the superior and inferior temporal gyri and in the right hemisphere in the anterior inferior frontal lobe, the temporo-parietal region including the supramarginal and angular gyri, and the inferior temporal gyrus.

When the BOLD activity generated by the Falling Questions (FQ) was compared to Falling Statements (FS) (Fig. 3C), there were no significant differences apart from those in the sensorimotor cortex and adjacent parietal cortex and in the occipital lobe.

Finally, when the BOLD activity generated by RQ was compared to FQ (Fig. 3D), increased activity was observed for RQ in the left hemisphere in the inferior frontal gyrus, in the right hemisphere in the inferior and adjacent middle and orbital frontal gyri, and in the anterior cingulate sulcus and the adjacent medial superior frontal gyrus.

Region of interest analysis

We selected ROIs based on the view that we should examine BOLD signal changes in areas that have been associated with processing acoustic, phonological, and sentential semantic structures. Accordingly, ROIs included primary and secondary acoustic processing areas in the superior temporal plane (including Sylvian fissure, Heschl's gyrus, the insula, and the central sulcus of the insula), areas that have been shown to be involved in phonological processing (inferior frontal areas, superior temporal sulcus, supramarginal gyrus, and angular gyrus) and areas involved in semantic processing (mid and inferior temporal gyrus, temporal pole, and inferior frontal areas). We also included those areas likely to be active from the push button activity in the motor cortex.

BOLD signal response for each stimulus type relative to baseline was analyzed in an ANOVA with the factors of hemisphere (2), ROI (17), and stimulus type (3). There was a statistically significant main effect of hemisphere ($F(1, 1121) = 5.1$, $P < 0.05$), a statistically significant main effect of ROI ($F(16, 1121) = 13.8$, $P < 0.0001$), and a trend towards significance for stimulus type ($F(2, 1121) = 3.22$, $P = 0.0613$). However, these results were qualified by a significant interaction of hemisphere, region, and stimulus type ($F(32, 1121) = 1.9$, $P < 0.005$). Tukey's test revealed significant differences in 9 regions, which are summarized in Table II. Bar graphs of the percentage signal change for each stimulus type in these 9 regions are shown in Figure 4.

The results of the ROI analysis can be summarized as follows:

1. There was a significant increase in BOLD activity for RQ over both FQ and FS in a number of right-sided ROIs (three temporal regions: the superior temporal gyrus, Heschl's gyrus, the superior temporal sulcus; and two inferior frontal regions: pars orbitalis and pars opercularis) and in one left-sided region (the superior temporal gyrus).
2. There were a number of ROIs on both sides that demonstrated an increase in BOLD signal for RQ over FQ only. Two were on the right (the circular sulcus of the insula and the inferior frontal sulcus) and three were on the left (Heschl's gyrus, inferior frontal gyrus, pars opercularis and triangularis).
3. One region, the pars triangularis in the left inferior frontal gyrus, demonstrated a trend towards increased activity for FS over FQ.

Table III summarizes the results of both the voxelwise and ROI analyses, omitting areas in the motor strip and visual cortex that were activated.

DISCUSSION

We created auditory stimuli consisting of triads of sentences that were identical in terms of their lexical content and propositional meaning and that differed in their intonational contour and illocutionary force. By comparing the vascular response generated in the brain by these stimuli, we can begin to identify brain regions that are responsible for the processing of intonation changes that underlie illocutionary force. Before discussing these implications of the study, we first consider several preliminary issues.

TABLE II. Areas of activation in the ROI analysis*

Region of interest	LH	P value	RH	P value
Sup temporal gyrus	RQ > FQ	<0.0001	RQ > FQ	<0.00001
	RQ > FS	0.0013	RQ > FS	0.065**
Sup temporal sulcus			RQ > FQ	0.08**
Heschl's gyrus	RQ > FS	0.014		
	RQ > FQ	<0.0001	RQ > FQ	<0.0001
	RQ > FS	<0.0001		
Sylvian fissure				
Middle temporal gyrus			RQ > FS	0.0001
Inf temporal gyrus				
Inf temporal sulcus				
Temporal pole				
Insula				
Circular sulcus of insula			RQ > FQ	0.0156
IFG-pars orbitalis			RQ > FQ	<0.0001
			RQ > FS	<0.0001
IFG-pars triangularis	RQ > FQ	<0.0001	RQ > FQ	<0.0001
			RQ > FS	0.0012
	FS > FQ	0.072**		
IFG-pars opercularis	RQ > FQ	0.023		
Inf frontal sulcus			RQ > FQ	0.0006
Supramarginal gyrus				
Angular gyrus				
Pre-central gyrus	FQ > FS	0.0003		

* Using Tukey's test for multiple comparisons (significance level $P < 0.05$), only those areas where significant differences between stimuli were found are reported. RQ = rising question; FQ = falling question; FS = falling statement; Sup = superior; Inf = inferior; IFG = inferior frontal gyrus.
** Trend towards significance.

First, the accuracy data demonstrate that the subjects discriminated these prosodic and word order contrasts during continuous echo-planar imaging. It has been shown that vascular responses to linguistic stimuli such as phonemes can be detected without modulation of scanner noise but that the response in the auditory cortex is somewhat blunted [Shah et al., 1999]. Work recently done on the effect of scanner noise has suggested that attention to external noise masking by the development of helmets and head wraps such as the one used in this study as well as attention to noise buffering within the MRI room itself improves auditory discrimination within the scanner [Ravicz and Melcher, 2001; Ravicz et al., 2000]. The accuracy rates in this study show that, with these techniques, subjects can make the contrasts we required despite scanner noise.

Second, subjects rated the different types of stimuli as equally emotionally intense, ruling out the possibility that differences in emotional intensity account for BOLD signal effects in this study. We must also consider the possibility that differences in emotional valence across the stimuli were confounded with illocutionary force; that is, that questions were perceived as more negative or positive than statements. This seems unlikely, but, even if it is true, we can nonetheless focus on the intonational determinants of illocutionary force. By comparing RQ and FQ stimuli, we examine the effect of rising intonation in determining question illocutionary force in questions only, thereby controlling for any possible differences in emotional valence between questions and statements.

Third, we must consider a possible strategy that subjects may have adopted in this study. All the FQ stimuli began with the words "was she," and the subjects might have made their judgements on the basis of the order of these initial words rather than on the basis of their assigning the illocutionary force of a question to the stimulus itself. This possibility is based upon the subtle distinction between a subject recognizing that the sentence-initial auxiliary-NP sequence indicates that a stimulus with an initial auxiliary is a question—the process that we tried to provoke—and a subject responding to a sentence-initial auxiliary-NP sequence with the "question" response "strategically"; i.e., because s/he inductively generalized over the first examples of this sort in the stimulus set to reach the conclusion that all stimuli with initial auxiliary-NP sequences would be questions in our study without actually computing illocutionary force in these stimuli.

There are three arguments against the view that subjects' "question" responses to these stimuli were generated by such a strategy without assigning their actual question illocutionary force. The first is that linguistic stimuli are processed at all levels of language when they are attended to. A listener who attends to an utterance cannot fail to assign some illocutionary force to that utterance. This is true even for the truncated and incomplete utterances that are so

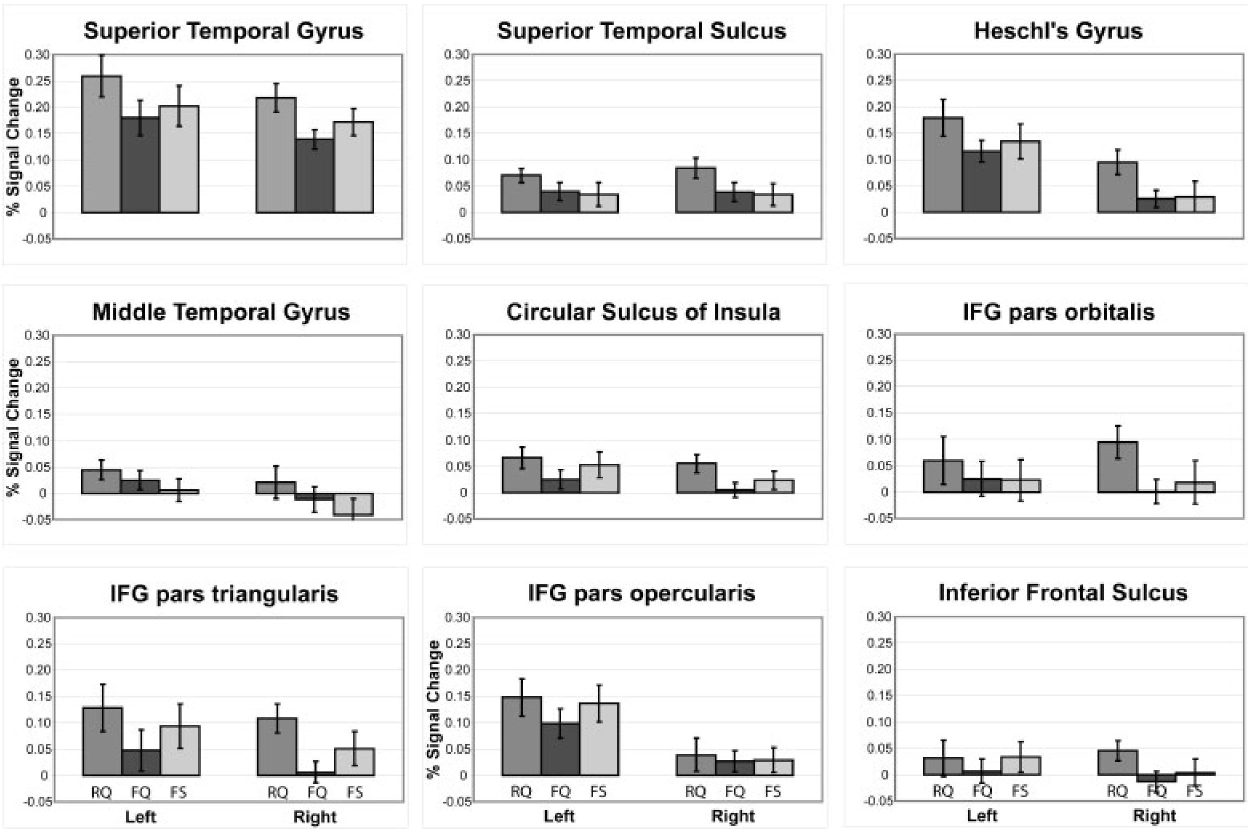

Figure 4.
Bar graphs of the mean percentage change [and standard error (SE)] in BOLD signal in the nine regions of interest where significant differences between the conditions were found. IFG = inferior frontal gyrus.

common in real language use. It is, therefore, very unlikely that subjects in this study responded with "question" to sentences with initial auxiliary-NP sequence entirely strategically, without actually assigning this illocutionary force to these utterances. Second, the fact that utterances with initial auxiliary-NP sequences were always questions is not just a

TABLE III. Summary of areas of activation*

	Voxelwise analysis		ROI analysis	
Effect	LH	RH	LH	RH
RQ > FS	STG, ITG	SMG/AG, ITG, IFG/frontal pole IFGorb/OG	STG	HG, STG, STS, MTG, IFG
FS > RQ				
FQ > FS				
FS > FQ			IFG	
RQ > FQ	IFG	Ant CS, IFG, IFG, IFGorb/OG	HG, STG, IFG	HG, STG, STS, IFG, IFS, CIS
FQ > RQ				

* AG, angular gyrus; CIS, circular sulcus of insula; CS, cingulate sulcus; HG, Heschl's gyrus; IFG, inferior frontal gyrus; IFGorb, inferior frontal gyrus pars orbitalis; IFS, inferior frontal sulcus; ITG, inferior temporal gyrus; MTG, middle temporal gyrus; OG, orbital gyrus; SMG, supramarginal gyrus; STG, superior temporal gyrus; STS, superior temporal sulcus.

fact about these utterances in our study, it is a fact of English: a sentence that begins with an auxiliary-NP sequence must be a question in English. If one thinks of the subjects' responses to these stimuli as a result of inductive generalization over instances of such stimuli that they were exposed to, one has to consider that such exposure was not limited to the first instances of these stimuli in our study but consisted of millions of subject-auxiliary inversion questions the subjects had heard before they were scanned. The result of such inductive generalization is not a strategy but an automatized process, the very process that underlies this comprehension. The situation in our study is, therefore, quite different from that in studies in which strategic performance occurs because subjects become sensitive to features of the stimulus set that are unique to the experiment. An anonymous reviewer also pointed out that the similarity in error rates for all stimulus types argues against the use of a strategy, which would be expected to lead to near-perfect performance. For these reasons, we would argue that the subjects in this study almost certainly assigned illocutionary force to these stimuli, and did not make their responses entirely through a strategic mapping of the initial words in these stimuli to a "question" response. They may well have made the assignment of question illocutionary force upon hearing the first words of these stimuli, but this too would be a normal process as utterances that begin with these sequences are always questions in English.

Fourth, as an anonymous reviewer pointed out, the intonational difference between stimuli in our study with rising and falling intonational contours was not limited to a rising final boundary tone in RQ stimuli. We focused on the rising final boundary tone in our description of these stimuli because a rising final boundary tone is the intonational feature that has been thought to be most important in turning an utterance that has the syntactic form of a statement into a question [Pierrehumbert, 1980]. As shown in Figure 1, this feature was present in the RQ stimuli and only these stimuli. However, as Figure 1 also illustrates, the FS F_o contour begins with a rise and then falls and the FQ F_o contour begins low, rises, and then falls. Subjects may have used some of these other intonational features to determine whether an utterance was a question or a statement, in one of two ways. First, though less important in determining illocutionary force than direction of the final boundary tone, these intonational features may also signal illocutionary values and could have been used by participants to determine these values. This would not vitiate the implications of this study, but rather make its focus the variety of intonational cues to illocutionary force found in natural utterances instead of a more narrow manipulation of the single dominant cue to illocutionary force. This seems appropriate in an initial study; more specific manipulation of particular intonational features is more appropriate in follow-up studies that could be undertaken once the areas related to processing intonation are initially delineated. Second, subjects may have learned that certain intonational features other than a rising or falling final boundary tone were reliably associated with questions and statements in this study, even if they are not universally associated with questions or statements. As with the role of strategic factors in determining responses to stimuli with subject-auxiliary inversion, the possibility that such strategic factors were the sole determinants of either the behavioral responses or BOLD signal effects seems remote. For expository convenience, we will continue to refer to the materials as being characterized by rising or falling intonation, recognizing that the differences across stimulus types are more complex than this nomenclature might imply.

The same reviewer pointed out that we did not include a fourth stimulus type: questions that were marked by both subject-auxiliary inversion and a rising intonation contour. Comparing such stimuli and FQ stimuli would add a contrast relevant to the localization of processing intonation in relationship to illocutionary force, and comparing RQ utterances to such stimuli would be useful in investigating the role of intonation in determining a subtle aspect of the meaning of RQ utterances (see below). However, the brain areas involved in using intonation to determine question illocutionary force can begin to be identified on the basis of the comparisons used in this study, as will be discussed below, even though we did not use these additional stimuli in this initial study.

With these preliminaries, we turn to a discussion of the differences in BOLD signal across conditions, seen in Table III. Our focus is on the brain regions involved in processing intonational contours with respect to illocutionary force. There is, however, one other finding that is worth mentioning first.

That finding is that comparisons of statements with falling intonation (FS) with questions with falling intonation (FQ), where intonation contour is held constant, showed increases in BOLD signal for statements, not questions, in the left IFG (in the ROI analysis). This activation may have resulted from subjects assigning the question illocutionary force to FQ stimuli earlier than the statement illocutionary force to FS stimuli (see discussion above). If this is not the case, however, this result is of interest because, intuitively, one might think of questions as being formed from their corresponding declarative propositions. Indeed, early versions of generative grammar theory explicitly incorporated this idea [Chomsky, 1957]. On the assumption that increases in BOLD signal reflect increased processing load or additional processing steps, the data here suggest that this is not the case. Two possibilities remain. One intuitively extremely unlikely possibility is that statements are formed from the corresponding questions. The more likely possibility is that neither statements nor questions are formed from the other, but that speakers and listeners assign illocutionary force as a discourse-level feature that is attached to the propositional content of a sentence. The FS>FQ effects in the ROI analysis provide evidence that the attribution of statement illocutionary force recruits part of the left perisylvian cortical region (the inferior frontal region) to a greater extent than the attribution of question illocutionary force. This localization is consistent with the fact that this is a classical language area of the brain in right handers. The data also provide evidence regarding the areas that are activated to a greater extent in association with the attribution of question illocutionary force than in association with the attribution of statement illocutionary force. This evidence is derived from both the RQ>FS and the RQ>FQ effects. We take the former to reflect differences in both illocutionary force and intonational contour and the latter to reflect a difference in intonational contour only. On this view, an area that is activated in the first but not the second comparison could reflect the attribution of question illocutionary force. There are two such areas, left and right anterior inferior temporal gyri, seen in the voxelwise analysis. The anterior ventral temporal cortex has been recognized as playing a role in the processing of meaning [e.g., Damasio et al., 1996; Halgren et al., 2002, Mazoyer et al., 1993; McCarthy et al., 1995; Nobre and McCarthy, 1995; Rossell et al., 2003; Smith et al., 1986; Vandenberghe et al., 2002]. Interestingly, these studies, which have focused on nominal concepts and propositional meaning, have generally found exclusively left hemisphere effects or bilateral effects that are significantly larger in the left hemisphere, while the effects of illocutionary force found in the present study were bilateral but larger over the right hemisphere.

Turning finally to the brain regions involved in processing intonational contours with respect to illocutionary force, the first point to note is that the comparisons of utterances with rising and falling intonation (RQ vs. FS and RQ vs. FQ) always showed increases in BOLD signal in the condition with rising intonation, never the reverse. Several aspects of processing a rising intonation contour could have produced these increases in BOLD signal.

One possibility is that they result from processing illocutionary value of a question at the semantic level. This, however, is ruled out because the effect of a rising intonation is seen in the comparison of questions with rising intonation (RQ) with questions with falling intonation (FQ), where question illocutionary force is held constant.

We note that, while the semantic processing of question illocutionary force is not likely to be the feature that is responsible for the increased BOLD signal found in RQ utterances than in either FS or FQ utterances, a more subtle illocutionary feature that is present only in RQ utterances, known as "conduciveness," may be responsible for this result. Conduciveness is the term given to questions where the answer is already known and the question is asked almost in disbelief or for emphasis, and is a feature of questions with the syntactic form of a statement and a rising intonation contour [Couper-Kuhlen, 1986; Glenn, 1977]. Thus, the increase in BOLD signal associated with RQ utterances may be due to a subtle illocutionary feature of RQ utterances. Note that, if the BOLD effect is due to the *assignment* of this semantic feature on the basis of the rising intonational contour in an utterance with the syntactic form of a statement, as opposed to the *presence* of the feature in the semantic representation of the utterance, the results speak to an aspect of the process of using intonation to assign an aspect of illocutionary force. In this case, the issue is whether that aspect is the assignment of question status or of conduciveness.

Another possible source of the increased BOLD signal in RQ utterances is processing the acoustic features of the rising intonation contour. Aspects of acoustic processing such as pitch discrimination in tones and phonemes, direction of pitch change in pure tones, and detection of spectral elements of speech sounds have been related to areas that were activated in the RQ>FQ and RQ>FS comparisons in this study [Belin et al., 2000; Bilecen et al., 1998; Burton et al., 2000; Demonet et al., 1992; Fiez et al., 1995; Johnsrude et al., 2000; Patel and Balaban, 2001; Paulesu et al., 1993; Sergent et al., 1992; Warren and Griffiths, 2003; Zatorre and Belin, 2001; Zatorre et al., 1992, 1994]. Thus, it is possible that the activation associated with utterances with a rising boundary tone is due to processing of the rising contour at the acoustic level.

Finally, it is also possible that the increased BOLD signal associated with RQ compared to FQ utterances reflects the process of interpreting the RQ intonation contour as a question. At a minimum, the regions in which BOLD signal increased in the RQ>FQ analysis can be thought to be a superset of the areas in which the process of interpreting the RQ intonation contour as a question takes place (this statement is, of course, subject to the limitations imposed on the detection of these areas related to the fMRI techniques, behavioral methods, and experimental design employed in this study). This study provides evidence that these areas are located in the perisylvian association cortex associated with other aspects of language processing, in both hemispheres, with what appears to be right-hemisphere predominance. The bilaterality of the BOLD signal effect is consistent with lesion studies, which have shown deficits of speech prosody involved in question/statement judgments in both left and right hemisphere–damaged patients [Baum et al., 1982; Blumstein and Goodglass, 1972; Behrens, 1988; Goodglass and Calderon, 1977; Heilman et al., 1975, 1984; Pell and Baum, 1997; Shapiro and Danly, 1985].

As we have said at several points, this is an early study of intonation, in which it was hoped to identify areas that are candidates for mapping of intonational contours onto illocutionary values. We conclude with an indication of how we think some of the issues that are not resolved by this preliminary study could be addressed.

The issue of whether the increased BOLD signal found in the RQ>FQ analysis reflects purely acoustic processing or mapping of intonational contours onto illocutionary values could be studied by measuring the response to utterances in which intonation is altered within and across illocutionary category boundaries. It has been shown that, if utterance-final pitch is increased in a step-wise fashion, listeners perceive the resulting utterance as a statement up to a certain point, and then quite abruptly perceive it as a question [Remijsen and van Heuven, 1999]. This phenomenon is similar to the well-known categorical perception effect found in perceiving many aspects of phonemes, and it provides an opportunity to disentangle purely acoustic from mapping determinants of BOLD signal effects in studies of intonation and illocutionary force. BOLD signal increases that are responsive to acoustic differences that cross categorical illocutionary boundaries and not to ones of identical size that are within illocutionary categories would suggest a phonological linguistic role in processing intonation for the regions in which they occur, whereas BOLD signal increases that are associated with within-category acoustic differences are likely to reflect non-categorical acoustic processing only.

The issue of whether the increased BOLD signal found in the RQ>FQ analysis reflects the attribution of conduciveness to RQ utterances can be approached contrasting the fourth stimulus type mentioned above—questions that are marked by both subject-auxiliary inversion and a rising intonation contour—with the FQ and RQ utterances used here. Regions that show increased BOLD signal in the first of these contrasts are ones associated with assigning question illocutionary force to acoustic features. These are expected to be a subset of those identified in the RQ>FS analysis. In the second contrast, an increase in BOLD signal in the RQ utterances compared to this fourth stimulus type would signal regions in which conduciveness is assigned to questions that have the syntactic form of statements and a rising intonational contour.

In summary, we have identified areas in both hemispheres in inferior frontal and temporal regions in which BOLD signal increased when subjects made question judgments about utterances with rising intonational contours compared to when they made question or statement judgments about utterances with falling intonational contours. The differences may reflect acoustic processing, assigning the illocutionary force of a question to a rising intonation contour, or to the presence of a subtle aspect of illocutionary force (conduciveness) in the utterances with rising intonational contours. All of these types of operations and features of utterances are tied to processing intonational contours, but only the second is specific to the processing of these contours as linguistic objects. Despite their interpretive limitations, these results provide new information relevant to the question of the brain regions involved in assigning the illocutionary force of a question to a rising intonation contour. The areas we have identified can be taken as including those areas. Further experimentation can move towards resolving some of the questions that remain unsettled.

ACKNOWLEDGMENTS

We gratefully acknowledge Doug Greve for invaluable assistance in the fMRI data analysis, Anders Dale and Bruce Fischl for numerous helpful discussions, Gina Kuperberg and Fujiro Ozawa for providing the parcellations used for the anatomical regions of interest, Evelina Busa for the cortical reconstructions, and Cody Evans and Sarah Groff for technical assistance.

REFERENCES

Austin J (1975): How we do things with words. Sbias M, Urmsson J, editors. Cambridge, MA: Harvard University Press.

Bandettini P (1993): MRI studies of brain activation: temporal characteristics. In: Functional MRI of the brain (workshop syllabus). Berkeley, CA: Society of Magnetic Resonance in Medicine; p 143–151.

Baum S, Daniloff R, Daniloff J, Lewis J (1982): Sentence comprehension by Broca's aphasics: effects of some suprasegmental variables. Brain Lang 17:261–271.

Behrens SJ (1988): The role of the right hemisphere in the production of linguistic stress. Brain Lang 33:104–127.

Belin P, Zatorre R, Lafaille P, Ahad P, Pike B (2000): Voice-selective areas in human auditory cortex. Nature 403:309–312.

Bilecen D, Scheffler K, Schmid N, Tschopp K, Seelig J (1998): Tonotopic organization of the human auditory cortex as detected by BOLD-FMRI. Hear Res 126:19–27.

Binder JR, Rao S, Hammeke T, Yeltkin F, Jesmanowicz A, Bandettini P, Wong E, Estkowski L, Goldstein M, Houghton R. (1994): Functional magnetic resonance imaging of human auditory cortex. Ann Neurol 35:662–672.

Binder J, Frost J, Hammeke S, Rao S, Cox R (1996): Function of the left planum temporale in auditory and linguistic processing. Brain 119:1239–1247.

Binder J, Frost J, Cox R, Rao S, Prieto T (1997): Human brain language areas identified by functional magnetic resonance imaging. J Neurosci 17:353–362.

Blumstein S, Cooper W (1974): Hemispheric processing of intonation contours. Cortex 10:146–158.

Blumstein S, Goodglass H (1972): The perception of stress as a semantic cue in aphasia. J Speech Hear Res 15:800–806.

Buchanan T, Lutz K, Mirzazade S, Specht K, Shah N, Zilles K, Jancke L (2000): Recognition of emotional prosody and verbal components of spoken language: an fMRI study. Cogn Brain Res 9:227–238.

Burock M, Dale A (2000): Estimation and detection of event related fMRI signals with temporally correlated noise: a statistically efficient and unbiased approach. Hum Brain Mapp 11:249–260

Burock M, Buckner R, Woldorff M, Rosen B, Dale A (1998): Randomized event-related experimental designs allow for extremely rapid presentation rates using fMRI. Neuroreport 9:3735–3739.

Burton M, Small S, Blumstein S (2000): The role of segmentation in phonological processing: an fMRI investigation. J Cogn Neurosci 12:679–690.

Caviness VS, Meyer J, Makris N, Kennedy D (1996): MRI-based topographic parcellation of human neocortex: an antomically specified method with estimate of reliability. J Cog Neurosci 8:566–587.

Chomsky N (1957): Syntactic structures. The Hague: Mouton.

Couper-Kuhlen E (1986): An introduction to English prosody. London: Edward Arnold.

Cox R (1996): AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comp Biomed Res 29:162–173.

Dale A (1999): Optimal experimental design for event-related fMRI. Hum Brain Mapp 8:109–114.

Dale A, Fischl B, Sereno M (1999): Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage 9:179–194.

Damasio H, Grabowski T, Tranel D, Hichwa R, Damasio A (1996): A neural basis for lexical retrieval. Nature 380:499–505.

Demonet J, Chollet F, Ramsay S, Cardebat D, Nespoulous J, Wise R, Rascol A, Frackowiak R (1992): The anatomy of phonological and semantic processing in normal subjects. Brain 115:1753–68.
Fiez J, Raichle M, Miezin F, Petersen S, Tallal P, Katz W (1995): Studies of auditory and phonological processing: Effects of stimulus characteristics and task demands. J Cogn Neurosci 7:357–375.
Fischl B, Sereno M, Dale A (1999a): Cortical surface-based analysis II Inflation, flattening, and surface-based coordinate system. Neuroimage 9:195–207.
Fischl B, Sereno MI, Tootell RBH, Dale AM (1999b): High-resolution inter-subject averaging and a coordinate system for the cortical surface. Hum Brain Mapp 8:272–284.
Fischl B, Liu A, Dale AM (2001): Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. IEEE Trans Med Imag 20:70–80.
Frith C, Friston K, Liddle P, Frackowiack R (1991): A PET study of word finding. Neuropsychologia 29:1137–1148.
Gabrieli J, Poldrack R, Desmond J (1998): The role of left pre-frontal cortex in language and memory. Proc Natl Acad Sci 95:906–913.
Gandour J, Wong D, Dzemidzic M, Lowe M, Tong Y, Li X (2003): A cross-linguistic fMRI study of perception of intonation and emotion in Chinese. Hum Brain Mapp 18:149–157.
George M, Parekh P, Rosinsky N, Ketter T, Kimbrell T, Heilman K, Herscovitch P, Post R (1996): Understanding emotional prosody activates right hemisphere regions. Arch Neurol 53:665–670.
Glenn M (1977): Pragmatic functions of intonation. PhD Thesis; University of Michigan, Ann Arbor.
Goodglass H, Calderon M (1977): Parallel processing of verbal and musical stimuli in right and left hemispheres. Neuropsychologia 15:397–407.
Halgren E, Dhond RP, Christensen N, Van Petten C, Marinkovic K, Lewine JD, Dale AM (2002): N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. Neuroimage 17:1101–1116.
Heilman K, Scholes R, Watson RT (1975): Auditory affective agnosia. JNNP 38:69–72.
Heilman K, Bowers D, Speedie L, Coslet H (1984): Comprehension of affective and non-affective prosody. Neurology 34:917–921.
Hickok, G. and D Poeppel (2000): Towards a functional neuroanatomy of speech perception. Trends Cogn Sci 4:131–138.
Howard D, Patterson K, Wise R, Brown W, Friston K, Weiller C, Frackowiak R (1992): The cortical localization of the lexicons. Brain 115:1769–1782.
Imaizumi S, Mori K, Kiritani S, Kawashima R, Sugiura M, Fukuda H, Itoh K, Kato T, Nakamura A, Hatano K, Kojima S, Nakamura K (1997): Vocal identification of speaker and emotion activates different brain regions. Neuroreport 8:2809–2812.
Johnsrude I, Penhune V, Zatorre R (2000): Functional specificity in the right human auditory cortex for perceiving pitch direction. Brain 123:155–163.
Kapur S, Rose R, Liddle P, Zipursky R, Brown G, Stuss D, Houle S, Tulving E (1994): The role of the left prefrontal cortex in verbal processing: semantic processing or willed action? Neuroreport 5:2193–2196.
Ladd D (1996): Intonational phonology. Cambridge: Cambridge University Press.
Mazoyer B, Tzourio-Mazoyer N, Frak V, Syrota A, Murayama N, Levrier O, Salamon G, Dehane S, Cohen L, Mehler J (1993): The cortical representation of speech. J Cogn Neurosci 5:467–479.
McCarthy G, Nobre AC, Bentin S, Spencer DD (1995): Language-related field potentials in the anterior-medial temporal lobe: I. Intracranial distribution and neural generators. J Neurosci 15: 1080–1089.
Meyer M, Alter K, Frieferici AD, Lohmann G, von Cramon YD (2002): FMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. Hum Brain Mapp 17:73–88.
Mitchell RL, Elliot R, Barry M, Cruttenden A, Woodruff PW (2003): Neural response to emotional prosody as revealed by functional magnetic resonance imaging. Neuropsychologia 41:1410–1421.
Nobre AC, McCarthy G (1995): Language-related field potentials in the anterior-medial temporal lobe: II. Effects of word type and semantic priming. J Neurosci 15:1090–1098.
Oldfield R (1971): The assessment and analysis of handedness: The Edinburgh Inventory. Neuropsychologia 9:97–113.
Pardo P, Makela J, Sams M (1999): Hemispheric differences in processing tone and amplitude modifications. Neuroreport 10: 3081–3086.
Patel A, Balaban E (2001): Human pitch perception is reflected in the timing of stimulus-related cortical activity. Nature Neurosci 4:839–844.
Paulesu E, Frith C, Frackowiak R (1993): The neural correlates of the verbal component of working memory. Nature 362:342–345.
Pell M, Baum S (1997): The ability to perceive and comprehend intonation in linguistic and affective contexts by brain damaged adults. Brain Lang 57:80–99.
Perani D, Dehaene S, Grassi F, Cohen L, Cappa SF, Dupoux E, Fazio F, Mehler J (1996): Brain processing of native and foreign languages. Neuroreport 7:2439–2444.
Pierrehumbert J (1980): The phonology and phonetics of English intonation. Cambridge, MA: MIT.
Poldrack R, Wagner A, Prull M, Desmond J, Glover G, Cabrieli J (1999): Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. Neuroimage 10:15–35.
Price C, Wise R, Warburton E, Moore C, Howard D, Patterson K, Frackowiak R, Friston K (1996): Hearing and saying: The functional neuroanatomy of hearing and saying. Brain 119:919–931.
Rademacher J, Galaburda A, et al (1992): Human cerebral cortex: Localization, parcellation and morphometry with magnetic resonance imaging. J Cogn Neurosci 4:352–374.
Ravicz M, Melcher J (2001): Isolating the auditory system from acoustic noise during functional magnetic resonance imaging: examination of noise conduction through the ear canal, head, and body. J Acoust Soc Am 109:216–231.
Ravicz M, Melcher J, Kiang N (2000): Acoustic noise during functional magnetic resonance imaging. J Acoust Soc Am, 108:1683–1696.
Remijsen B, van Heuven VJ (1999): Gradient and categorical pitch dimensions in Dutch: Diagnostic test. Proceedings of the 11th International Congress of the Phonetic Sciences, San Francisco 3:1865–1868.
Ross E (1981): The Aprosodias: functional anatomic organization of the affective components of language in the right hemisphere. Arch Neurol 38:561–569.
Ross E, Mesulam M (1979): Dominant language functions in the right hemisphere: Prosody and emotional gesturing. Arch Neurol 36:144–148.
Ross E, Thompson R, Yenkosky J (1997): Lateralization of affective prosody in brain and the callosal integration of hemispheric language functions. Brain Lang 56:27–54.
Rossell SL, Price CJ, Nobre AC (2003): The anatomy and time course of semantic priming investigated by fMRI and ERPs. Neuropsychologia 41:550–564.

Schlanger B, Schlanger P, Gerstmann L (1976): The perception of emotionally toned sentences by right hemisphere damaged and aphasic subjects. Brain Lang 3:396–403.
Searle J (1969): An essay in the philosophy of language. Cambridge, MA: Cambridge University Press.
Selkirk E (1995): Sentence prosody: intonation, stress and phrasing. In: Goldsmith J, editor. The handbook of phonological theory. Boston: Blackwell; p 551–567.
Sergent J, Zuck E, Levesque M, MacDonald B (1992): Positron emission tomography study of letter and object processing: empirical findings and methodological considerations. Cereb Cortex 2:68–80.
Shah N, Jancke L, Grosse-Ruyken M, Muller-Gartner H (1999): Influence of acoustic masking noise in fMRI of auditory cortex during phonetic discrimination. J Magn Reson Imag 9:19–25.
Shapiro B, Danly M (1985): The role of the right hemisphere in the control of speech prosody in propositional and affective contexts. Brain Lang 25:19–36.
Smith ME, Stapleton JM, Halgren E (1986): Human medial temporal lobe potentials evoked in memory and language tasks. Electroencephalogr Clin Neurophysiol 63:145–159.
Stiller D, Gaschler-Markefski B, Baumgart F, Schindler F, Tempelman C, Heinze H, Scheich H (1997): Lateralized processing of speech prosodies in the temporal cortex: A 3-T functional magnetic resonance imaging study. MAGMA 5:275–284.
Talairach J, Tournoux P (1988): Co-planar stereotactic atlas of the human brain: 3-dimensional system: an approach to cerebral imaging. Stuttgart: Thieme.
Tallal P, Miller S, Fitch R (1993): Neurobiological basis of speech: a case for the preeminence of temporal processing. Ann NY Acad Sci 682:27–47.
Tucker D, Watson R, Heilman K (1977): Discrimination and evocation of affectively intoned speech in patients with right parietal disease. Neurology 27:947–950.
Turner R, Howseman A, Rees G, Josephs O (1997): Functional imaging with magnetic resonance. In: Frackowiak F, Frith CD, Dolan RJ, Mazziotta JC, editors. Human brain functions. New York: Academic Press; p 467–486.
Van Lancker D, Sidtis J (1992): The identification of affective prosodic stimuli by right and left hemisphere damaged subjects: all errors are not equal. J Speech Hear Res 35:963–970.
Vandenberghe R, Nobre AC, Price CJ (2002): The response of the left temporal cortex to sentences. J Cogn Neurosci 14:550–560.
Vandenberghe R, Price C, Wise R, Josephs O, Frackowiack R (1996): Functional anatomy of a common semantic system for words and pictures. Nature 383:254–256.
Warburton E, Wise R, Price C, Weiller C, Hadar U, Ramsay S, Frackowiak RS (1996): Noun and verb retrieval by normal subjects. Brain 119:159–179.
Warren JD, Griffiths TD (2003): Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. J Neurosci 23:5799–5804.
Wildgruber D, Pihan H, Ackermann H, Erb M, Grodd W (2002): Dynamic brain activation during processing of emotional intonation: influence of acoustic parameters emotional valence and sex. Neuroimage 15:856–869.
Wise R, Chollet F, Hadar U, Friston K, Hoffner E, Frackowiack R (1991): Distribution of cortical neural networks involved in word comprehension and word retrieval. Brain 114:1803–1817.
Wolfe G, Ross E (1987): Sensory aprosodia with left hemiparesis from subcortical infarction. Right hemisphere analogue of sensory-type aphasia with right hemiparesis? Arch Neurol 44:661–671.
Zatorre R, Belin P (2001): Spectral and temporal processing in human auditory cortex. Cereb Cortex 11:946–953.
Zatorre R, Evans A, Meyer E, Gjedde A (1992): Lateralization of phonetic and pitch discrimination in speech processing. Science 256:846–849.
Zatorre R, Evans A, Meyer E (1994): Neural mechanisms underlying melodic perception and memory for pitch. J Neurosci 14:1908–1919.
Zurif E, Mendelsohn M (1972): Hemispheric specialization for speech sounds: The influence of intonation and structure. Percept Psychophysic 11:329–332.