FMRI Reveals Brain Regions Mediating Slow Prosodic Modulations in Spoken Sentences

Martin Meyer,[1,2]* Kai Alter,[1] Angela D. Friederici,[1] Gabriele Lohmann,[1] and D. Yves von Cramon[1]

[1]*Max-Planck-Institute of Cognitive Neuroscience, Leipzig, Germany*
[2]*Centre for Functional Imaging Studies, Edinburgh, Scotland*

Abstract: By means of fMRI measurements, the present study identifies brain regions in left and right peri-sylvian areas that subserve grammatical or prosodic processing. Normal volunteers heard 1) normal sentences; 2) so-called syntactic sentences comprising syntactic, but no lexical-semantic information; and 3) manipulated speech signals comprising only prosodic information, i.e., speech melody. For all conditions, significant blood oxygenation signals were recorded from the supratemporal plane bilaterally. Left hemisphere areas that surround Heschl gyrus responded more strongly during the two sentence conditions than to speech melody. This finding suggests that the anterior and posterior portions of the superior temporal region (STR) support lexical-semantic and syntactic aspects of sentence processing. In contrast, the right superior temporal region, in especially the planum temporale, responded more strongly to speech melody. Significant brain activation in the fronto-opercular cortices was observed when participants heard pseudo sentences and was strongest during the speech melody condition. In contrast, the fronto-opercular area is not prominently involved in listening to normal sentences. Thus, the functional activation in fronto-opercular regions increases as the grammatical information available in the sentence decreases. Generally, brain responses to speech melody were stronger in right than left hemisphere sites, suggesting a particular role of right cortical areas in the processing of slow prosodic modulations. *Hum. Brain Mapping 17:73–88, 2002.* © **2002 Wiley-Liss, Inc.**

Key words: speech perception; speech prosody; intonation; modulation; functional magnetic resonance imaging (fMRI); hemispheric lateralization; insula; superior temporal region; frontal operculum

INTRODUCTION

In auditory language processing distinct peri-sylvian areas serve phonological properties [Binder et al., 1994; Jäncke et al., 1998; Wise et al., 1991], enable access to the meaning of words [Démonet et al., 1992; Fiez et al., 1996], or process the structural relations between words [Caplan et al., 1998; Dapretto and Bookheimer, 1999; Friederici et al., 2000a; Mazoyer et al., 1993]. Thus, a broadly distributed network involving the entire peri-sylvian cortex in the left hemisphere plays a predominant role in word-level, linguistic processing [Binder et al., 1996; Démonet et al., 1994].

Only a relatively small number of brain imaging studies have investigated fluent speech at the sen-

tence-level. From these studies an inconsistent picture has emerged. Listening to normal sentences occurs automatically, thereby recruiting the superior temporal region bilaterally, and relegating the inferior frontal cortex to a minor role [Dehaene et al., 1997; Friederici et al., 2000a; Kuperberg et al., 2000; Meyer et al., 2000; Müller et al., 1997; Schlosser et al., 1998]. Studies explicitly focusing on the processing of syntactic information during sentence comprehension, however, report selective activation in left fronto-lateral and fronto-opercular cortices [Caplan et al., 1998; Dapretto and Bookheimer, 1999; Friederici et al., 2000a].

Finding right cortical areas engaged during speech comprehension, either at the word-, or even more prominently, at the sentence level, challenges the notion that auditory language processing is an exclusively left-hemisphere operation [Corina et al., 1992; Goodglass, 1993; Zurif, 1974]. Given prior imaging data and traditional views, a re-investigation is needed concerning the functional role of the right superior temporal region in the analysis of aurally presented fluent speech.

Prosodic information in spoken language comprehension

An adequate description of the cerebral organization of auditory sentence comprehension must consider the role of prosodic information. Following a universally accepted definition, prosody can be described as "the organizational structure of speech" [Beckman, 1996]. One purpose of this paper is therefore to investigate the processing of one prosodic attribute: relatively slow F_0 movements that define speech melody (intonation). In intonation languages such as German, French, and English, fundamental frequency (F_0) together with other prosodic characteristics such as duration and intensity, are used to convey 1) information about affective and emotional states and 2) the locations of intonational phrase boundaries and phrasal stresses [Shattuck-Hufnagel and Turk, 1996]. The occurrence of F_0 markers at intonational phrase boundaries serves an important linguistic function, because the boundaries of intonational phrases often coincide with the boundaries of syntactic constituents. Evidence from a recent ERP study indicates that intonational boundaries are immediately decoded and used to guide the listener's initial decisions about a sentence's syntactic structure [Steinhauer et al., 1999; Steinhauer and Friederici, 2001]. Many studies have shown that attention to prosodic cues (i.e., F_0 and duration) can help listeners to distinguish between potentially ambiguous sentences,

e.g., old (*men and women*) vs. (*old men*) and women [Baum and Pell, 1999; Price et al., 1991]. The types of slow F_0 movements that occur at phrasal boundaries can also give listeners some semantic information about the sentences they are listening to. For example, speakers use different F_0 movements to signal questions vs. statements, where questions are usually produced with an F_0 rise at the end of the utterance. Thus, slow F_0 movements can extend over chunks of utterances longer than just one segment and therefore constitute the speech melody. These relatively slow pitch movements contrast with rapid pitch movements that sometimes can be observed in the vicinity of consonantal constrictions.

Intonation, e.g., slow pitch movement, serves a variety of linguistic functions and is therefore a major prosodic device in speech. Because prosody provides linguistic information that is not available from orthographic transcriptions of utterances, speech comprehension therefore requires rapid decoding of both grammatical and prosodic parameters to achieve a final interpretation of a spoken utterance.

This raises the question whether the neural substrates of prosodic processing are part of the cerebral network subserving speech or if processing prosodic information involves distinct areas in the human brain.

Although there is strong evidence that processing affective intonation is lateralized to the right hemisphere [Blonder et al., 1991; Bowers et al., 1987; Darby, 1993; Denes et al., 1984; Ross et al., 1997; Starkstein et al., 1994] the functional lateralization of linguistically relevant intonation remains a subject of debate [Baum and Pell, 1999; Pell, 1998; Pell and Baum, 1997]. One hypothesis proposes that right cortical areas subserve the linguistic aspects of prosody [Brådvik et al., 1991; Bryan, 1989; Pell and Baum, 1997; Weintraub and Mesulam, 1981]. The competing claim suggests that the processing of linguistic aspects of prosody is represented in the left hemisphere whereas the right hemisphere subserves the processing of emotional prosody [Baum et al., 1982; Emmorey, 1987]. An alternative explanation proposes that functional lateralization of prosodic cues correlates with distinct physical parameters of prosody: a left hemisphere dominance for duration and amplitude with the latter encoding loudness at the perception level and a right hemisphere dominance for F_0 as an acoustic correlate of speech melody [Van Lancker and Sidtis, 1992]. The distinction between emotional and linguistic prosody cannot be made until the physical correlates that distinguish these two types of prosodic phenomena have been adequately described [Pell, 1998].

Evidence from a brain lesion study implies, in fact, that left as well as right cortical areas are selectively involved when unique prosodic attributes have to be processed [Behrens, 1985]. Thus, the left hemisphere appears essential in processing stress contrasts, whereas the right hemisphere predominantly engages in processing slow F_0 movement. This was the case for low-pass-filtered sentences, which lack grammatical information, and sound like speech but are almost unintelligible. Further evidence was provided by a dichotic listening study when the intonation contour of an unintelligible low-pass-filtered sentence had to be compared to a preceding one [Blumstein and Cooper, 1974]. Finding a consistent left ear superiority for these tasks suggested greater right hemisphere involvement in the processing of intonation contour, i.e., slow pitch movements.

On the basis of the combined results, it could be hypothesized that right peri-sylvian areas are more strongly involved when speech melody has to be processed. Left hemisphere areas, however, come into play more strongly when sentence prosody is combined with matching grammatical information, be it lexical meaning or syntactic structure.

Present study

The present study investigated the hemodynamic correlates of speech melody perception at the sentence level. Three sentence conditions varied in lexical and syntactic information. In the *normal speech* condition syntactic, lexical, and prosodic information was available. In the *syntactic speech* condition, lexical information was omitted by replacing all content words with phonotactically legal pseudowords whereas syntactic information (functional elements such as determiners, auxiliaries, and verb inflections) remained unaltered. In the *prosodic speech* condition all lexical and syntactic information had been removed yielding a speech utterance reduced to speech melody. Listening to delexicalized speech melody allows no access to word forms or morpho-syntactic information.

Based on earlier studies we predicted the comprehension of normal speech would engage the STR bilaterally, with the interior frontal gyrus (IFG) being activated to a weaker extent. In a number of recent brain imaging studies the auditory processing of normal sentences occurred quite automatically in the STR bilaterally, with only marginal or no activation in Broca's area [Dehaene et al., 1997; Kuperberg et al., 2000; Meyer et al., 2000; Müller et al., 1997; Schlosser et al., 1998]. In a recent fMRI study the bilateral inferior frontal cortex, however, subserved processing of syntactic speech, focusing on syntactic and prosodic information in the absence of lexical information [Friederici et al., 2000a]. Pronounced left inferior frontal activation also corresponds to the processing of syntactically complex sentences thereby involving aspects of syntactic memory [Caplan et al., 1999]. These data suggest that auditory comprehension of less complex sentences requires only minor recruitment of the left inferior frontal cortex. According to these findings, the salience of inferior frontal responses increases directly with the level of syntactic processing. Thus, the inferior frontal cortex, particularly in the left hemisphere, is expected to participate in the processing of syntactic speech.

Temporal cortices are known to be involved in general syntactic functions, i.e., morpho-syntactic processing at the sentence level [Friederici et al., 2000a; Humphries et al., 2001; Meyer et al., 2000; Stowe et al., 1998]. Thus, this region ought to respond less during the prosodic speech condition that completely lacks syntactic and lexical information. Furthermore we predict to find decreased responses to prosodic speech in the auditory cortex bilaterally because the total amount of auditory information contained in prosodic speech is clearly reduced relative to normal and syntactic speech. According to recent PET-studies on pitch perception [Gandour et al., 2000; Zatorre et al., 1992, 1994] a bilateral engagement of the inferior frontal cortex is also likely. The data at hand, however, allow no clear prediction with respect to the lateralization of the frontal activation: although processing pitch variations in a grammatically relevant context was shown to occur in the left fronto-opercular cortex [Gandour et al., 2000], pitch processing of complex auditory stimuli in a nonlinguistic context rather revealed either right dorsolateral activation [Zatorre et al., 1992] or bilateral contributions of the frontal operculum [Zatorre et al., 1994]. By presenting speech melody, we will be able to test which hemisphere responds more sensitively to language-related slow prosodic modulations. Taken together, it is likely that cortical areas in the right hemisphere will be more strongly involved while subjects listen to prosodic as compared to normal or syntactic speech.

MATERIALS AND METHODS

Subjects

Fourteen right-handed subjects (8 male, mean age $\pm$ SD, 25.2 $\pm$6 years) participated in the study after giving written informed consent in accordance with the guidelines approved by the Ethics Committee of

the Leipzig University Medical Faculty. Subjects had no hearing or neurological disorders and normal structural MRI scans. They had no prior experience with the task and were not familiar with the stimulus material.

Stimuli

The corpus comprised three experimental conditions at the sentence-level.

Normal speech

Normal speech consists of grammatically, semantically, and phonologically correct sentences (18 active and 18 passive mode readings).

Active Mode:
Die besorgte Mutter sucht das weinende Kind.
The anxious mother searches for the crying child.
Passive Mode:
Das weinende Kind wird von der besorgten Mutter gesucht.
The crying child is searched for by the anxious mother.

Syntactic speech

Syntactic speech consists of grammatically correct pseudo sentences (18 active and 18 passive mode readings), i.e., with correct syntax and phonology. Note that sentences contain phonotactically legal pseudo words in the place of content words. Morphological inflections, determiners, and auxiliaries, however, remain unaltered, so that syntactic assignments are still possible with the sentence meaning removed.

Active Mode:
Das mumpfige Folofel hongert das apoldige Trekon.
The mumpfy folofel hongers the apolding trekon.
Passive Mode:
Das mumpfige Folofel wird vom apoldigen Trekon gehongert.
The mumpfy folofel is hongered by the apolding trekon.

Prosodic speech

Prosodic speech consists of grammatically uninterpretable sentences, that contain only suprasegmental information, i.e., prosodic parameters such as global F_0 contour and amplitude envelope. To achieve an acoustic signal that is exclusively reduced to its suprasegmental information normal speech files were treated by applying the PURR-filtering procedure [Sonntag and Portele, 1998] (see Figures 5 and 6). The signal derived from this filtering procedure does not comprise any segmental or lexical information.

From a linguistic point of view, F_0 contour and amplitude envelope represent speech melody—the distribution and type of pitch accents and boundary markers of prosodic domains. Prosodic speech could be described as sounding like speech melody listened to from behind a door.

All sentences of Condition 1 and 2 were recorded with a trained female speaker in a sound proof room (IAC) at a 16 bit/41.1 kHz sampling rate, then digitized and downsampled at a 16 bit/16 kHz sampling rate. All stimuli except delexicalized signals were normalized in amplitude (70%). Because the latter were limited in bandwidth compared to the other two conditions a stronger normalization (85%) was necessary to guarantee equal intensity. The mean length of the sentences in the normal speech condition was ±SD, 3.59 ± 0.35 sec, in the syntactic speech condition ±SD, 3.76 ± 0.37 sec, and in the prosodic speech condition ±SD, 3.59 ± 0.35 sec.

Procedure

Participants were presented with 108 stimuli, 36 for each condition, in a pseudo-random order. The stimulus corpus comprised of only unique sentences: sentences were not repeated during the experiment. The sounds were presented binaurally through headphones designed specifically for use in the scanner. A combination of external ear defenders and perforated ear plugs that conducted the sound directly into the auditory passage were used to attenuate the scanner noise without affecting the quality of speech stimulation. Subjects were asked to judge via pressing a button immediately after presentation of each sentence (either normal, syntactic or prosodic speech) whether the stimulus was realized in active or in passive tense. The design assigned a fixed presentation rate. Each trial started with a warning tone (1,000 Hz, 200 msec) 1,500 msec before the speech input. The epoch-related approach was chosen for the purposes of the present study. To enable an epoch-related data analysis successive presentation of single sentences was separated by an Inter-Trial-Interval of rest lasting 10 sec. Thus, the delayed hemodynamic response is allowed to return to a baseline level.

Data acquisition

MRI data were collected at 3.0 T using a Bruker 30/100 Medspec system (Bruker Medizintechnik

GMBH, Ettlingen, Germany). The standard bird cage head coil was used. Before MRI data acquisition, field homogeneity was adjusted by means of "global shimming" for each subject. Then, scout spin echo sagittal scans were collected to define the anterior and posterior commissures on a midline sagittal section. For each subject, structural and functional (echo-planar) images were obtained from eight axial slices parallel to the plane intersecting the anterior and posterior commissures (AC–PC plane). The most inferior slice was positioned below the AC–PC plane and the remaining seven slices extended dorsally. The whole range of slices comprised an anatomical volume of 46 mm, covered all parts of the peri-sylvian cortex, and extended dorsally to the intraparietal sulcus. After defining the slices' position conventional T1 weighted anatomic images (MDEFT: TE 10 msec, TR 1,300 msec, in plane-resolution 0.325 mm^2) were collected in plane with the echo-planar images, to align the functional images to the 3D-images [Norris, 2000; Ugurbil et al., 1993]. A gradient-echo EPI sequence was used with a TE 30 msec, flip angle 90°, TR 2,000 msec, acquisition bandwidth 100 kHz. Acquisition of the slices within the TR was arranged so that the slices were all rapidly acquired followed by a period of no acquisition to complete the TR. The matrix acquired was 64 × 64 with a FOV of 19.2 cm, resulting in an in-plane resolution of 3 mm^2. The slice thickness was 4 mm with an interslice gap of 2 mm. In a separate session high resolution whole-head 3D MDEFT brain scans (128 sagittal slices, 1.5 mm thickness, FOV 25.0 × 25.0 × 19.2 cm, data matrix of 256 × 256 voxels) were acquired additionally for reasons of improved localization [Lee et al., 1995; Ugurbil et al., 1993].

Data analysis

The fMRI data were processed using the software package *LIPSIA* [Lohmann et al., 2001]. The first trial of each experimental block was excluded from analysis to avoid influence of vascular arousal caused by the onset of the scanner noise. During reconstruction of the functional data the corresponding runs were concatenated into a single run. Data preparation included correction for movements, i.e., the images of the fMRI time series were geometrically aligned using a matching metric based on linear correlation. Here, timestep 50 served as reference. To correct for the temporal offset between functional slices acquired in one image, a sinc-interpolation algorithm based on the Nyquist-Shannon Theorem was employed. Low frequency signals were suppressed by applying a 1/32 Hz high-pass filter (2 times the length of one complete

oscillation). The increased autocorrelation caused by temporal filtering was taken into account during statistical evaluation by the adjustment of the degrees of freedom.

To align the functional dataslices onto a 3D stereotactic coordinate reference system, a rigid linear registration with six degrees of freedom (3 rotational, 3 translational) was calculated yielding an individual transformation matrix. The rotational and translational parameters were acquired on the basis of 2D-MDEFT slices to achieve an optimal match between individual 2D structural data and 3D reference data set. Secondly, each individual transformation matrix was scaled to the standard Talairach brain size (x = 135, y = 175, z = 120) [Talairach and Tournoux, 1988] by applying a linear scaling. The resulting parameters were then used to align the 2D functional slices with the stereotactic coordinate system by means of a trilinear interpolation, generating datasets with a spatial resolution of 3 mm^3. Additionally, this linear normalization was improved by a subsequent non-linear normalization [Thirion, 1998].

The statistical analysis was based on a least squares estimation using the general linear model (GLM) for serially autocorrelated observations [Aguirre et al., 1997; Bosch, 2000; Friston, 1994; Zarahn et al., 1997]. For each individual subject, statistical parametric maps (SPM) were generated with the standard hemodynamic response function and a response delay of 6 sec. The model equation including the observation data, the design matrix, and the error term, was linearly smoothed by convolving it with a Gaussian kernel of dispersion of 4 sec FWHM. The contrast between the different conditions was calculated using the t-statistics. Subsequently, t-values were converted to Z-scores. As the individual functional datasets were all aligned to the same stereotactic reference space a group analysis of fMRI-data was performed by averaging individual Z-maps. The average SPM was multiplied by a SPM correction factor of the square root of the current number of subjects (n = 14) [Bosch, 2000]. For the purpose of illustration, averaged data were superimposed onto one normalized 3D MDEFT standard volume. Averaged SPMs were thresholded with $|Z| \geq 3.1$ ($P < 10^{-3}$ one-tailed, uncorrected for multiple comparisons).

Region of interest analyses

To test whether brain responses to prosodic speech are more strongly lateralized to right peri-sylvian areas as compared to the remaining sentence conditions statistical testing by means of spherical regions of

TABLE I. Behavioral data*

Condition	Correctness (%)	Standard error (±)
Normal speech	93.35	2.77
Syntactic speech	91.76	2.90
Prosodic speech	51.78	1.92

* Grammatical judgment performance revealed approximately perfect correctness for normal and syntactic speech, but chance rate performance for prosodic speech.

interest (ROIs) was performed [Bosch, 2000]. Functional activation of each condition was calculated separately. In this comparison, the Inter-Stimulus-Interval (resting period) after each single trial presentation served as the baseline for data analysis. To compare condition-specific brain activation with respect to functional lateralization, spherical ROIs were defined bilaterally yielding three temporal ROIs (anterior, mid, and posterior STR) within each hemisphere. All spherical ROIs contained the local maximal Z-score of the summed functional activation of all conditions within a 4 mm radius (see Table V).

In a second step normalized Z-scores were averaged within 3 bilateral spherical ROIs of each participant yielding mean Z-scores for each condition [Bosch, 2000]. The mean Z-scores were subjected to repeated-measures ANOVA to analyze differences in local brain activation across subjects and between conditions. All main effects of interactions with two or more degrees of freedom in the numerator were adjusted

with the procedure suggested by Huynh and Feldt [1970].

RESULTS

Behavioral data

As apparent from Table I, subjects responded correctly consistently on over 90% of normal and syntactic speech sentences in the grammatical task. Thus, the sentences were quite easily understood. In contrast, response to prosodic speech was at chance level clearly demonstrating that the applied filtering procedure left no grammatical information available.

fMRI data

Three main contrasts were performed, the first to reveal regions corresponding to syntactic speech (normal vs. syntactic speech), the second and third to reveal brain regions subserving prosodic speech (normal vs. prosodic speech) and (syntactic vs. prosodic speech). Inter-subject averaging revealed significant peri-sylvian activation bilaterally that varied systematically as a function of intelligibility. The results are presented in Tables II–IV and Figure 1.

Normal vs. syntactic speech

In both the left and right hemisphere normal speech activated the posterior temporo-occipital junction

TABLE II. Normal speech vs. syntactic speech*

Location	BA	Left hemisphere				Right hemisphere			
		Z score	x	y	z	Z score	x	y	z
Normal speech > syntactic speech									
POTA	39	4.53	−46	−56	25	4.59	37	−64	26
Syntactic speech > normal speech									
IFS		—	—	—	—	−4.42	37	11	23
IPCS		—	—	—	—	−4.26	36	31	15
FOP		−4.2	−34	26	5	—	—	—	—
ROP	44/6	—	—	—	—	−5.77	43	8	−2
STR	22/41/42	−6.77	−50	−6	2	−7.23	44	−25	11

* This table and Tables III and IV list results of averaged Z maps based on individual contrasts between conditions. To assess the significance of an activation focus averaged Z maps were thresholded with $|Z| \geq 3.1$ ($P < 10^{-3}$ one tailed, uncorrected for multiple comparisons). Localization is based on stereotactic coordinates (Talairach and Tournoux, 1988). These coordinates refer to the location of maximal activation indicated by the Z score in a particular anatomical structure. Distances are relative to the intercommissural (AC–PC) line in the horizontal (x), anterior-posterior (y) and vertical (z) directions. The table only lists activation clusters exceeding a minimal size of 150 voxels. IFG, inferior frontal gyrus; IFS, inferior frontal sulcus; MFG, middle frontal gyrus; IPCS, inferior precentral sulcus; aINS, anterior insula; FOP, deep frontal operculum; ROP, Rolandic operculum; STR, superior temporal region; POTA, parieto-occipital transition area; aCG, anterior cingulate gyrus; pCG, posterior cingulate gyrus; PT, planum temporale; PP, planum parietale; Tha, thalamus; Cau, caudate head; Put, putamen.

TABLE III. Normal speech vs. prosodic speech[*]

Location	BA	Left hemisphere				Right hemisphere			
		Z score	x	y	z	Z score	x	y	z
Normal speech > prosodic speech									
IFG	45	8.3	−49	25	13	—	—	—	—
STR	22/41/42	11.1	−46	−20	6	10.73	49	−8	0
pCG	30	6.83	−9	−52	20	—	—	—	—
POTA	39	5.98	−43	−57	22	—	—	—	—
Tha	—	4.9	−15	−28	2	—	—	—	—
Prosodic speech > normal speech									
MFG	10/46	−6.29	−33	43	10	—	—	—	—
aINS	—	−7.04	−34	17	7	−6.5	27	20	3
ROP	44/6	—	—	—	—	−7.7	43	10	8
PT/PP	42/22	—	—	—	—	−6.49	49	−32	20
mCG	23	−5.34	−5	−19	30	—	—	—	—
Cau	—	−6.5	−15	18	6	−5.04	10	16	7

* Functional activation indicated separately for contrasts between conditions. For explanation see Table II.

area. Hemodynamic responses to syntactic speech were greater in the STR bilaterally, in the right precentral sulcus and Rolandic operculum. Furthermore, frontal activation foci in the right inferior frontal sulcus and in the left deep frontal operculum were found.

Normal vs. prosodic speech

As apparent from Table III local blood supply to the STR bilaterally was significantly greater in the normal speech condition. In addition, a stronger signal was observed in a particular segment of the left IFG (pars triangularis), in the left posterior cingulate gyrus, in the area lining the posterior MTG, and in the angular gyrus. Additionally, the thalamus was more strongly

involved when subjects listened to normal speech. The delexicalized prosodic speech activated brain regions in the anterior half of the insula and in the deep frontal operculum bilaterally. Also, the right Rolandic operculum, the right posterior Sylvian fissure, the left middle frontal gyrus, the left middle cingulate gyrus as well as the caudate head bilaterally were more strongly recruited by the pure speech melody.

Syntactic vs. prosodic speech

Like normal speech the syntactic speech condition also corresponds to stronger brain activation in the STR bilaterally as well as in the left IFG (pars triangularis). Furthermore, several subcortical activation foci

TABLE IV. Syntactic speech vs. prosodic speech[*]

Location	BA	Left hemisphere				Right hemisphere			
		Z score	x	y	z	Z score	x	y	z
Syntactic speech > prosodic speech									
IFG	45	9.53	−49	26	19	—	—	—	—
STR	22/41/42	16.02	−45	−20	6	14.82	48	−10	3
Put	—	5.87	−26	−7	−1	4.19	11	−8	−4
Tha	—	4.89	−6	−14	8	—	—	—	—
Tha	—	5.87	−14	−29	3	—	—	—	—
Prosodic speech > syntactic speech									
MFG	10/46	−4.7	−33	43	11	—	—	—	—
aINS	—	−5.21	−34	17	7	−4.44	26	19	4
ROP	44/6	—	—	—	—	−4.84	41	10	12
PT/PP	42/22	—	—	—	—	−4.83	54	−39	22
Cau	—	−5.26	−13	17	6	—	—	—	—

* Functional activation indicated separately for contrasts between conditions. For explanation see Table II.
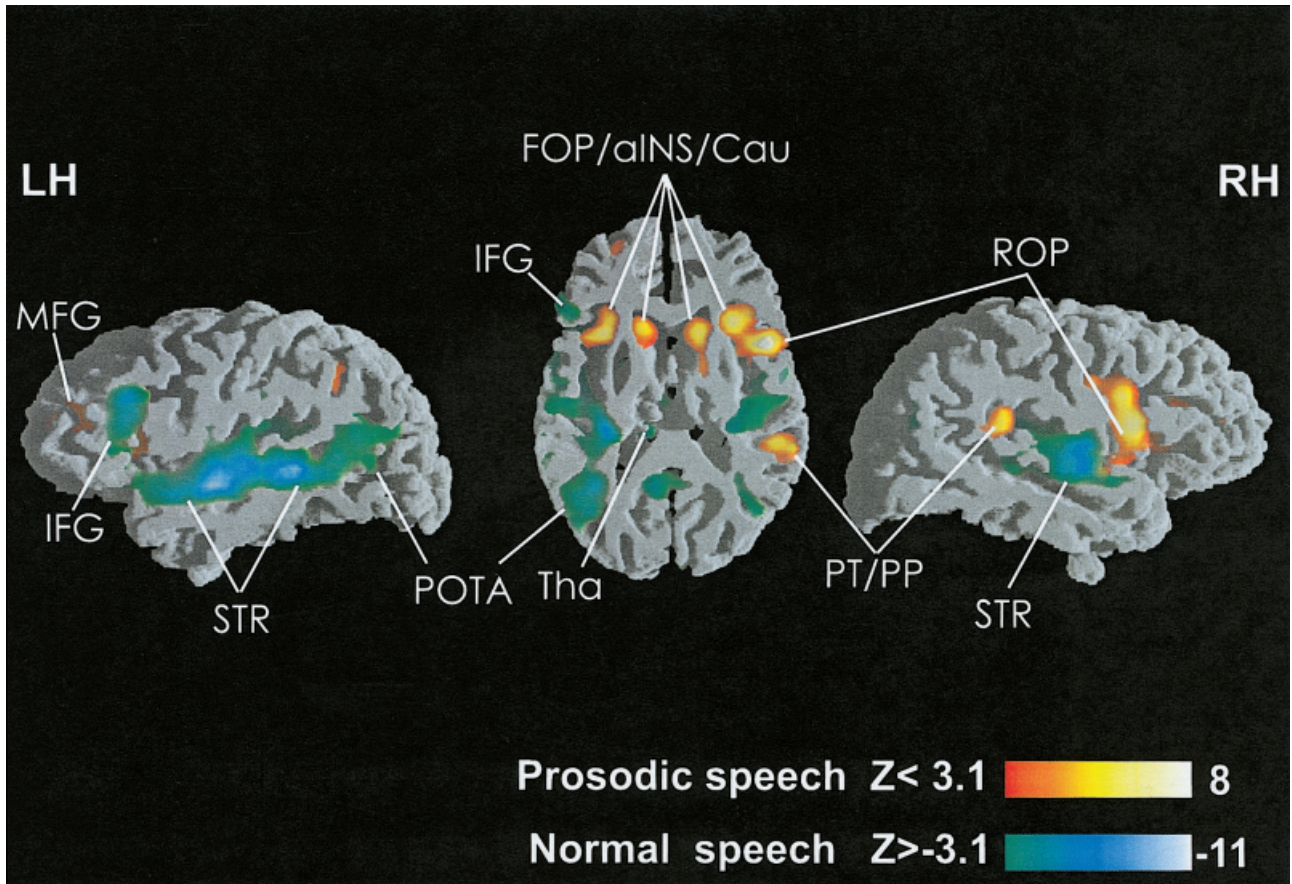
Figure 1.

Views of direct comparison between normal and prosodic speech. Functional inter-subject activation exceeding the significance threshold is depicted on parasagittal and horizontal slices intersecting the peri-sylvian cortex ($n = 14$). In all images hemodynamic responses to normal speech are illustrated by means of the greenish color scale, activation for prosodic speech is illustrated by means of the reddish color scale.
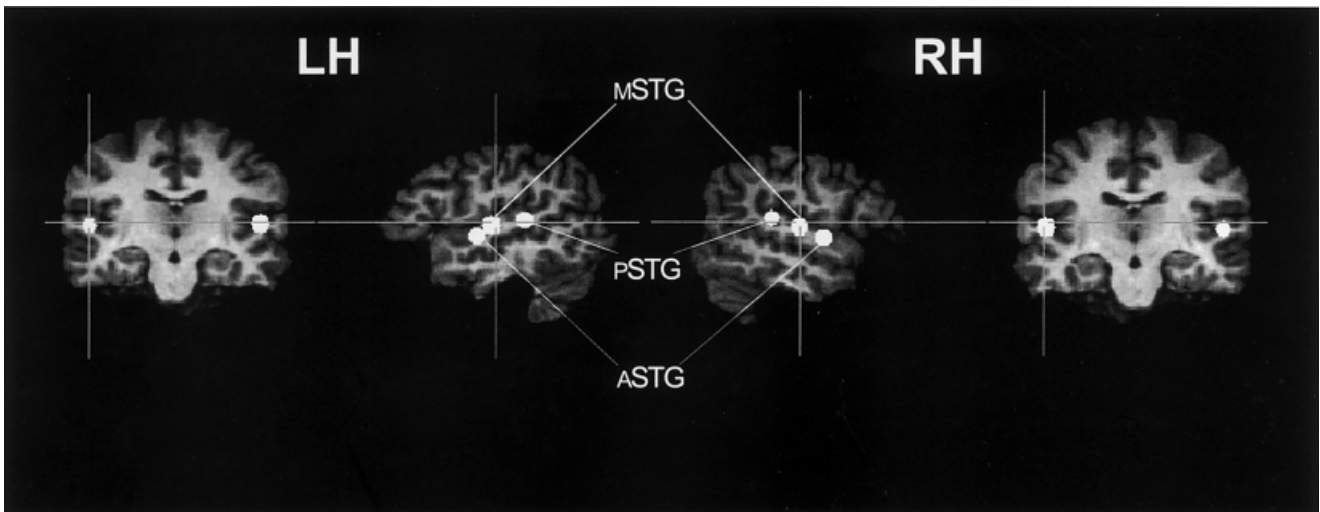

Figure 2.

Location of fixed spherical ROIs placed in both the right (RH) and the left (LH) supratemporal plane (see also Table V).

(basal ganglia, thalamus) were significantly activated. Listening to prosodic speech produced stronger activation in the anterior half of the insula and the deep frontal operculum bilaterally, the right Rolandic operculum, the right posterior Sylvian fissure, and the subcortical caudate head bilaterally. Thus, the processing of unintelligible speech corresponds to reduced activation in the STR bilaterally when compared to intelligible utterances. In contrast, the left and right fronto-opercular cortex was most strongly activated for prosodic speech. In addition, data presented in Figure 1 indicate that hemodynamic responses to normal speech are stronger in the left as compared to the right hemisphere, whereas areas involved more strongly in prosodic speech appear to be larger in the right as compared to the left hemisphere. The results of a further statistical analysis using spherical regions of interest (ROIs) addresses this issue exclusively.

Statistical analysis of ROIs

Figure 3 indicates mean Z-scores for experimental conditions obtained from six bilateral spherical ROIs [Bosch, 2000] whose positions are shown in Figure 2.

Irrespective of the condition inter-subject activation revealed significant superior temporal activation bilaterally with the left STR revealing reduced spatial extent of functional brain responses to prosodic speech (cf. Fig. 4). Irrespective of the hemisphere, prosodic speech produces the weakest activation when compared to normal and syntactic speech, whereas brain responses are strongest for syntactic speech. As predicted, the increase in brain activation was more strongly lateralized to right cortical sites for prosodic speech when compared to normal and syntactic speech.

For the purpose of testing statistical differences in local brain activation, mean Z-scores of spherical ROIs were subjected to systematic analyses (ANOVA). As predicted different hemodynamic responses in particular brain regions (factors ROI, hemisphere) varied as a function of the specific aspects of normal speech, syntactic speech, and prosodic speech (factor condition).

Global effects

A global ($2 \times 3 \times 3$) ANOVA with factors hemisphere $\times$ ROI $\times$ condition revealed a main effect of condition ($F_{2,26} = 34.08$, $P < 0.0001$), a main effect of ROI ($F_{2,26} = 20.60$, $P < 0.0001$), and an interaction
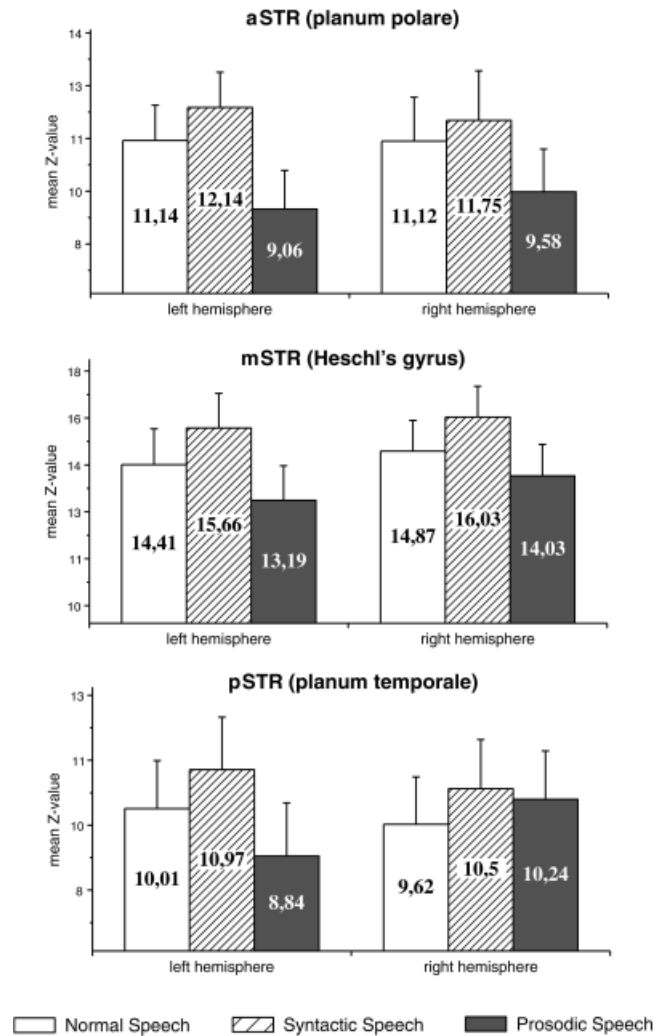

Figure 3.
Mean Z-scores obtained from 6 spherical regions of interest (ROIs). Error bars refer to the SE.

between ROI $\times$ condition ($F_{4,52} = 7.29$, $P < 0.005$) indicating a different pattern of brain responses in distinct regions of interest. No main effect of hemisphere was found, but two interactions hemisphere $\times$ condition ($F_{2,26} = 12.89$, $P < 0.0005$), and hemisphere $\times$ ROI $\times$ condition ($F_{4,52} = 3.41$, $P < 0.05$) show that activation varied within hemispheres as a function of different conditions.

Right superior temporal region and prosodic speech

Given the data apparent from Figure 3 the result obtained from the global ANOVA can be explained by the consistent right hemisphere superiority of brain responses to prosodic speech relative to normal and syntactic speech. Consequently, a ($2 \times 3 \times 2$) ANOVA
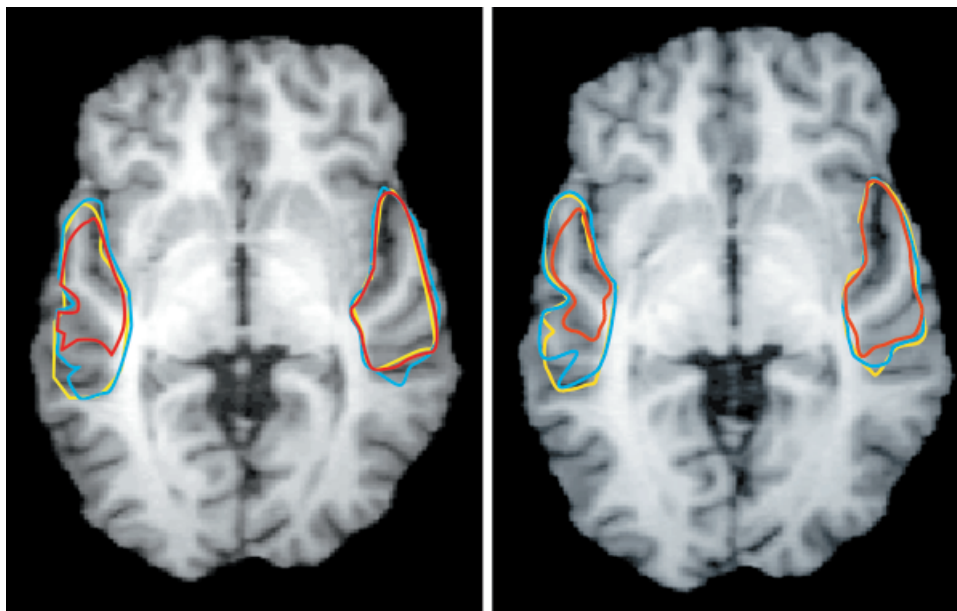
Figure 4.

Superimposing of functional brain responses obtained from the STR ($Z \geq 15$). Cortical volume activated by normal and syntactic speech is marked with yellow and blue lines. The red line indicates the extension of the activation cluster elicited by prosodic speech. Functional inter-subject activation on contiguous horizontal slices reveals the sensitivity of the left supratemporal plane for grammatical, that is lexical and syntactic information, whereas no differences could be found in the contralateral cortex.

with factors hemisphere × ROI × condition for normal and syntactic speech only revealed a main effect of condition ($F_{1,13} = 44.24$, $P < 0.0001$), but no interactions between factors ROI × condition, ROI × hemisphere, hemisphere × condition, and ROI × hemisphere × condition. Thus, neither normal nor syntactic speech affect one hemisphere more than the contralateral site.

Based on this finding, a new variable *segmental speech* was generated, by averaging the Z-scores for normal and syntactic speech within each single ROI. Segmental and prosodic speech were compared in an (2 × 3 × 2) ANOVA with factors hemisphere × ROI × condition. For this comparison a main effect of condition ($F_{1,13} = 31.83$, $P < 0.0001$), and several in-

teractions between factors ROI × condition ($F_{2,26} = 7.99$, $P < 0.005$), hemisphere × condition ($F_{1,13} = 16.42$, $P < 0.005$), and ROI × hemisphere × condition ($F_{2,26} = 4.33$, $P < 0.05$) were found, clearly indicating that right temporal sites are more strongly involved in processing prosodic speech when compared to normal and syntactic speech.

DISCUSSION

We discuss the brain activation data for temporal and frontal cortices and then detail the specific contribution of the right peri-sylvian cortex to prosodic processing.

TABLE V. Position of all spherical ROIs*

Region	Left			Right		
	x	y	z	x	y	z
aSTR (planum polare)	−50	−8	−1	47	−6	0
mSTR (Heschl's gyrus)	−49	−14	5	47	−17	7
pSTR (planum temporale)	−51	−32	12	52	−31	15

* Coordinates of centre voxels are listed in stereotactic space (Talairach and Tournoux, 1988). Each spherical ROI contains 257 voxels (radius=4mm).
aSTR, anterior superior temporal region; mSTR, middle superior temporal region; pSTR, posterior superior temporal region.
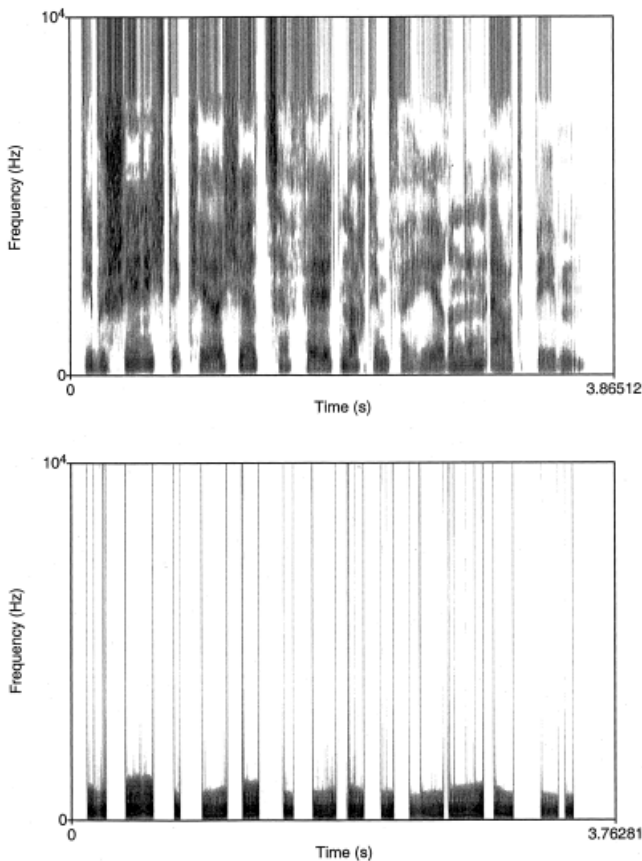
right hemisphere. This finding consistently replicates the activation pattern described by a previous fMRI study [Friederici et al., 2000a]. Normal and syntactic speech only differ with respect to the presence of lexical-semantic information, so that the activation difference emerging from this comparison could be explained by the absence of lexical-semantic information in the syntactic speech condition. As all meaningful words were replaced by phonotactically legal pseudowords the increase in superior temporal activation is taken to reflect additional processing when unknown pseudowords are heard and subjects failed to find an equivalent lexical entry. This observation is consistent with previous neuroimaging studies that
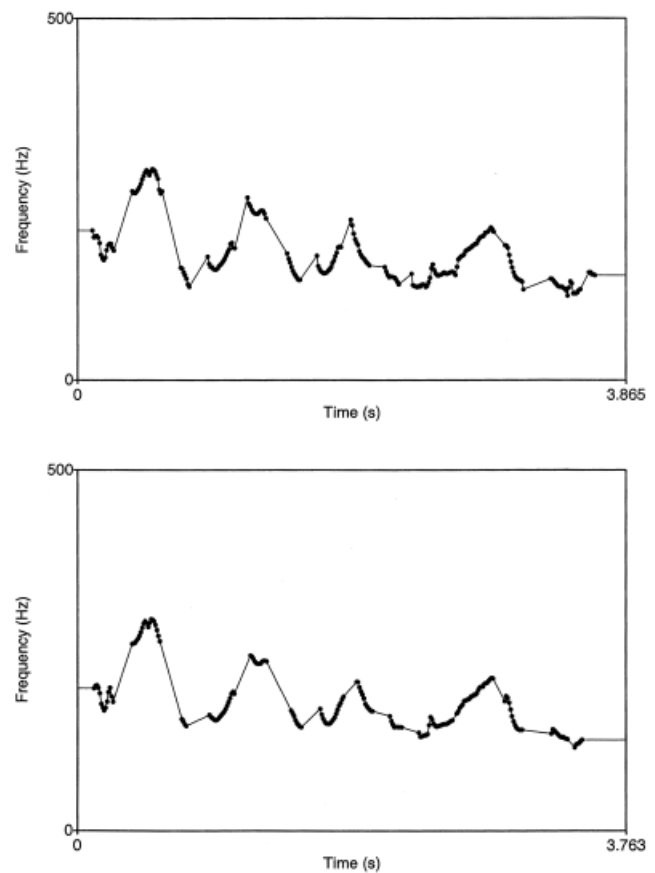

Figure 5.

Wide band spectrogram of speech signals before and after application of the PURR-filter. The upper spectrogram illustrates the frequency spectrum (0–10 KHz) of a normal sentence. The lower image illustrates clearly the reduced spectral information derived from a PURR-filter treated prosodic speech stimulus. The acoustic signal derived from this filtering procedure is reduced to frequencies containing the F_0 as well as the 2nd and 3rd harmonic. Additionally, all aperiodic portions of the speech signal are removed.

Temporal cortex

The present study identified peri-sylvian areas bilaterally subserving speech comprehension at the sentence level. This finding is in general accordance with recent neuroimaging studies on the processing of spoken words [Binder et al., 2000; Mazoyer et al., 1993] and sentences [Dehaene et al., 1997; Kuperberg et al., 2000; Meyer et al., 2000; Müller et al., 1997; Schlosser et al., 1998] that demonstrated activation in anterior and posterior divisions of the STR. In accordance with the predictions, however, the functional activation in the STR varied as a function of intelligibility.

Activation for syntactic speech as compared to normal speech was greater in the STR of both the left and

Figure 6.

Pitch contour of speech signals before and after application of the PURR-filter. The upper image illustrates the pitch contour of a normal sentence. Peaks and valleys of pitch contour symbolize typical sentence intonation. The lower image shows pitch contour for the same delexicalized sentence suggesting that the filtering procedure does not change the intonation contour. Duration of the delexicalized sentence is shorter as compared to the normal sentence because the aperiodic portions of the speech signal are removed.

showed that the perceptual analysis of normal speech signals is subserved by the temporal lobes [Zatorre et al., 1992]. A similar explanation is given by Hickok and Poeppel [2000], who also argue that the sound-based representation of speech is located in the posterior-superior temporal region bilaterally. With respect to this interpretation one objection can be raised. To detect whether a sentence was in the passive or active mode, subjects merely had to detect the presence of "wird" plus "das" or "von" plus "ge-t" at the end of a sentence. The actual lexical content of the content words in normal speech was irrelevant to the task. Thus, the task actually might not completely distinguish syntactic from normal speech. In a previous fMRI study, however, we also presented participants with the same set of normal and syntactic speech sentences [Friederici et al., 2000a]. In this study, subjects performed a different task. They had to decide whether the stimulus contained a syntactic structure and content words. To master this task subjects had to process function words and content words in the normal speech condition. In this recent study, brain activation in the STR bilaterally was also greater for syntactic speech when compared to normal speech. The combined data suggest that subjects process the actual lexical elements of a spoken utterance even if it is irrelevant to the task.

For prosodic speech, when compared to normal and syntactic speech, significantly weaker activation in both the left and the right STR was observed. This finding supports the predictions because delexicalized sentences do not contain phonological, lexical-semantic, or syntactic, but only prosodic information. This finding is also consistent with a recent fMRI study, which reported activation decrease in auditory cortices when subjects listened to incomprehensible speech [Poldrack et al., 2001]. Conversely, functional activation in the left and right STR is significantly increased for normal and syntactic speech indicating that grammatical processing during auditory sentence presentation is supported by the STR of both the left and the right hemisphere. This result is in general agreement with recent neuroimaging studies associating auditory sentence comprehension with an involvement of the bilateral STR with the left hemisphere playing a dominant role in right-handed subjects [Kuperberg et al., 2000; Müller et al., 1997; Schlosser et al., 1998]. In particular, the (left) anterior STR was shown to play an essential role in processing constituent structures, i.e., general syntactic operations at the sentence-level [Friederici et al., 2000b; Humphries et al., 2001; Meyer et al., 2000; Scott et al., 2000]. In line with these earlier studies, the present study reveals greater activation in the anterior and mid STR relative to posterior STR regions during sentence processing (cf. Fig. 3).

In terms of functional lateralization a different pattern can be described in particular, when considering the supratemporal plane (cf. Figs. 3,4). For the right STR a clear functional lateralization for prosodic speech was observed.

Processing pure pitch movements correspond to larger activation in the posterior part of the right STR. Further evidence for a special role for the right STR in pitch processing comes from two recent PET studies. Perry et al. [1999] compared a condition in which subjects repeatedly produced a single vocal tone of constant pitch to a control in which they heard similar tones at the same rate. Within the STG only right-sided activation were noted in primary auditory and adjacent posterior supratemporal cortices. The authors interpret this finding as a reflection of the specialization of the right auditory cortex for pitch perception because accurate pitch perception of one's own voice is required to adjust and maintain the intended pitch during singing. Tzourio et al. [1997] found a significant rightward asymmetry between Heschl gyri and the posteriorly situated planum temporale for passive listening to tones compared to rest. Based on their findings they propose that right supratemporal regions, in particular the planum temporale, play an elementary role in pitch processing. These findings are in agreement with the present data because prosodic speech also produces significantly more activation in the right posterior part of the Sylvian Fissure when compared to normal speech (c.f. Fig. 1). Thus, the right posterior STR is supposed to play a vital role in processing pitch information available in auditory stimuli.

Relative to acoustically (and hence grammatically) more complex normal and syntactic speech, processing prosodic speech involves a smaller volume in the left supratemporal plane (cf. Fig. 4). In contrast, processing speech stimuli recruits the left STR more strongly. Thus, the left STR is particularly sensitive to rapid pitch movements building speech rather than to slow pitch movements representing speech melody. A similar hierarchical organization of the human auditory cortex was recently displayed by an fMRI study: the processing of pure tones is confined to auditory core areas whereas belt areas support more complex sounds [Wessinger et al., 2001].

With respect to the right STR processing prosodic speech involves a large cortical volume surrounding the auditory core regions. This observation is in accord with prior lesion and imaging data that suggest a greater involvement of right rather than left supratem-

poral areas in certain aspects of pitch processing [Johnsrude et al., 2000; Zatorre and Belin, 2001; Zatorre and Samson, 1991]. Furthermore, a recent PET-study provided evidence for the existence of voice-selective regions in the human cortex [Belin et al., 2000]. According to these data the perception of voices involves left and right cortical regions in the STR. The perception of prosodic speech may be reminiscent of human vocal utterances and therefore recruits voice-sensitive areas in the superior temporal regions.

Frontal cortex

As predicted sentence processing does not significantly involve Broca's area. This finding agrees with several recent neuroimaging studies that also report only minor activation in the inferior frontal cortex during auditory sentence comprehension [Dehaene et al., 1997; Mazoyer et al., 1993; Meyer et al., 2000; Müller et al., 1997; Schlosser et al., 1998; Scott et al., 2000]. A selective engagement of left inferior cortex, i.e., Broca's area, is reported by studies that investigated the processing of syntactic information during auditory and written sentence comprehension thereby focussing on syntactic complexity [Caplan et al., 1998, 2000]. In the present experiment the syntactic speech condition also involved syntactic processing but focussed on a different aspect. Subjects merely had to process the morpho-syntactic information (auxiliaries and prefix ge-) to accomplish the task. Additionally, the stimulus corpus consisted of syntactically simple sentence structures that can be parsed quite easily. The activation focus for the syntactic speech condition is not located in the lateral crown of Broca's area, but in the directly adjacent-buried fronto-opercular cortex. This finding is consistent with a previous fMRI study that also presented subjects with syntactic speech [Friederici et al., 2000a]. The combined findings can be taken to show that the fronto-opercular region does not subserve the processing of syntactic complexity. Furthermore, a recent fMRI study has addressed the issue of syntactic complexity and its relation to Broca's area explicitly [Fiebach et al., 2001]. This study used German sentences that allowed the distinguishing of syntactic complexity and syntactic memory requirements. The results demonstrated that the function of Broca's area must be attributed to aspects of syntactic memory. Thus, it appears that Broca's area is involved in syntactic memory. This view is not inconsistent with the interpretation that Broca's area mediates verbal working memory [Paulesu et al., 1993; Pöppel, 1996].

Normal and syntactic speech activated an inferior frontal area in the left hemisphere that was localized anterior to Broca's area situated in the left IFG (pars triangularis). According to recent neuroimaging studies investigating language comprehension this region is associated with a semantic executive system whose function is to access, maintain and manipulate semantic information represented elsewhere in the cortex [Poldrack et al., 1999]. Both syntactic and normal speech contain function words, but only the latter contains content words representing lexical-semantic information. It might be possible that the left inferior fronto-lateral activation observed for normal and syntactic speech reflects the maintenance and manipulation of elements at a pre-lexical level, i.e., operations of syllabification, mandatory in both conditions.

Processing syntactic and prosodic speech corresponds to strong bilateral fronto-opercular hemodynamic responses. The amount of activation increases from syntactic speech to prosodic speech. Thus, the bilateral activation in the fronto-opercular cortex might be attributed to the unsuccessful effort of extracting syntactic and lexical-semantic information from degenerate stimuli. This observation is consistent with recent fMRI data also demonstrating that speech comprehension and related effort interact in fronto-opercular areas [Giraud et al., 2001]. Additionally, another recent fMRI study reported that the right inferior cortex responds stronger to incomprehensible relative to comprehensible speech [Poldrack et al., 2001].

According to these findings it is conceivable that the right fronto-opercular cortex seeks primarily to extract slow pitch information from the speech stream, whereas the left fronto-opercular region mainly performs the extraction of segmental units from the speech signal. This speculation is partially supported by a recent PET study reporting a portion of the Broca area involved in extraction of segmental units from lexical tones [Gandour et al., 2000].

The stronger right lateralized premotor activation in the Rolandic operculum found for prosodic speech when compared to normal and syntactic speech (cf. Tables III,IV) also needs to be clarified. This asymmetry is consistent with one fMRI study reporting right inferior precentral activation for both covert and overt singing with the contralateral response pattern emerging during both overt and covert speech production [Riecker et al., 2000]. We might speculate therefore that in the present study subjects reproduced silently the tune of delexicalized sentence melody to master the task instruction.

Right hemisphere superiority in prosodic speech

The finding that functional activation in right relative to left peri-sylvian areas in the prosodic speech condition indicates a functional right hemisphere predominance for processing speech melody. The present fMRI data obtained from normal subjects are in agreement with results from several lesion studies reporting a right hemisphere dominance for the perception of intonation contour [Heilman et al., 1984]. Right-brain-damaged (RBD) patients as well as left-brain-damaged (LBD) patients showed decreased comprehension of linguistic prosody (question vs. statement) relative to that of normal controls. Additionally, only RBD patients revealed deficient processing of emotional sounding speech. On the basis of these results the authors propose that a right hemisphere dysfunction cause global prosodic-melodic defects including the processing of speech melody. Another recent study investigated the prosodic competences of RBD and LBD patients. The results indicate that right hemisphere's involvement in prosody comprehension increased when the linguistic significance of speech was reduced [Perkins et al., 1996]. This finding is also corroborated by a lesion study investigating the lateralization of linguistic prosody in patients with either a left or right temporal lesion during sentence production [Schirmer et al., 2001]. Although LBD patients revealed more difficulties in timing their speech production, the RBD patients were mainly impaired in producing speech melody. An unusual F_0 production at the sentence level in a population of RBD patients was also observed by Behrens [1989]. Thus, it appears that the right hemisphere plays an essential role in analyzing various acoustic cues critical for the performance of non-linguistic, auditory tasks.

In summary, the present results are in accord with the view that linguistic prosody is a right hemisphere function [Brådvik et al., 1991; Bryan, 1989; Weintraub and Mesulam, 1981].

When considering the functional lateralization of linguistic prosody, however, a crucial question is to what extent different prosodic parameters can be isolated and located in each of the hemispheres as it was previously suggested by van Lancker and Sidtis [1992]. In the present study, a right hemisphere preponderance in processing intonation consisting of slow pitch movements has been demonstrated.

CONCLUSION

Our study demonstrated the differential involvement of human frontal and temporal cortex in auditory sentence processing varying as a function of different linguistic information types. Processing lexical and syntactic (grammatical) information predominantly involves the left superior temporal region. In contrast, processing prosodic parameters (speech melody) appears to involve the contralateral cortex.

Functional brain activation in left and right fronto-opercular cortices increased as a function of unintelligibility, i.e., when the proportion of linguistic information available in the speech stimuli decreases. Thus, the fronto-opercular cortices might provide additional neuronal resources whenever the brain seeks to combine syntactic and prosodic information to achieve a successful interpretation.

Brain responses to speech melody were generally stronger in right than left hemisphere sites, pointing to a higher susceptibility of right cortical areas for processing slow pitch movements.

ACKNOWLEDGMENTS

REFERENCES

Aguirre GK, Zarahn E, D'Esposito M (1997): Empirical analysis of BOLD-fMRI statistics. II. Spatially smoothed data collected under null-hypothesis and experimental conditions. Neuroimage 5:199–212.

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

Baum SR, Pell MD (1999): The neural basis of prosody: insights from lesion studies and neuroimaging. Aphasiology 13:581–608.

Beckman M (1996): The parsing of prosody. Lang Cognit Processes 11:17–68.

Behrens S (1985): The perception of stress and lateralization of prosody. Brain Lang 26:332–348.

Behrens S (1989): Characterizing sentence intonation in a right-hemisphere damaged population. Brain Lang 37:181–200.

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

Binder JR, Frost JA, Hammeke TA, Bellgowan PSF, Springer JA, Kaufman JN, Possing ET (2000): Human temporal lobe activation by speech and nonspeech sounds. Cereb Cortex 10:512–528.

Binder JR, Frost JA, Hammeke TA, Rao SM, Cox RW (1996): Function of the left planum temporale in auditory and linguistic processing. Brain 119:1239–1254.

Binder JR, Rao SM, Hammeke TA (1994): Functional magnetic resonance imaging of human auditory cortex. Ann Neurol 35:662–672.

Blonder LX, Bowers D, Heilman KM (1991): The role of the right hemisphere in emotional communication. Brain 114:1115–11127.

Blumstein S, Cooper WE (1974): Hemispheric processing of intonation contours. Cortex 10:147–158.

Bosch V (2000): Statistical analysis of multi-subject fMRI data: the assessment of focal activations. J Magn Reson Imaging 11:61–64.

Bowers D, Coslett HB, Bauer RM, Speedie LJ, Heilman KM (1987): Comprehension of emotional prosody following unilateral hemispheric lesions: processing defect vs. distraction effect. Neuropsychologia 25:317–328.

Brådvik B, Dravins C, Holtås S, Rosen I, Ryding E, Ingvar D (1991): Disturbances of speech prosody following right hemisphere infarcts. Acta Neurol Scand 84:114–126.

Bryan K (1989): Language prosody in the right hemisphere. Aphasiology 3:285–299.

Caplan D, Alpert N, Waters G (1998): Effects of syntactic structure and propositional number on patterns of regional cerebral blood flow. J Cog Neurosci 10:541–552.

Caplan D, Alpert N, Waters G (1999): PET studies of syntactic processing with auditory sentence presentation. Neuroimage 9:343–351.

Caplan D, Alpert N, Waters G, Olivieri A (2000): Activation of Broca area by syntactic processing under conditions of concurrent articulation. Hum Brain Mapp 9:65–71.

Corina D, Vaid J, Bellugi U (1992): The linguistic basis of left hemisphere specialization. Science 255:1258–1260.

Dapretto M, Bookheimer SY (1999): Form and content: dissociating syntax and semantics in sentence comprehension. Neuron 24:427–432.

Darby DG (1993): Sensory aprosodia: a clinical clue to the lesions of the inferior division of the right middle cerebral artery. Neurology 43:567–572.

Dehaene S, Dupoux E, Mehler J, Cohen L, Paulesu E, Perani D, van de Moortele P, Lehricy S, Le Bihan D (1997): Anatomical variability in the cortical representation of first and second language. Neuroreport 8:3809–3815.

Démonet JF, Chollet F, Ramsay S, Cardebat D, Nespoulous J, Wise R, Rascol A, Frackowiak RSJ (1992): The anatomy of phonological and semantic processing in normal subjects. Brain 115:1753–1768.

Démonet JF, Price CJ, Wise R, Frackowiak RSJ (1994): Differential activation of right and left posterior sylvian regions by semantic and phonological tasks: a positron-emission tomography study in normal human subjects. Neurosci Lett 182:25–28.

Denes G, Caldognetto EM, Semenza C, Vagges K, Zettin M (1984): Discrimination and identification of emotions in human voice by brain-damaged subjects. Acta Neurol Scand 69:154–162.

Emmorey K (1987): The neurological substrates for prosodic aspects of speech. Brain Lang 30:305–320.

Fiebach C, Schlesewsky M, Friederici A (2001): Syntactic working memory and the establishment of filler-gap dependencies: insight form ERPs and fMRI. J Psycholinguist Res 30:321–338.

Fiez JA, Raichle ME, Balota DA, Tallal P, Petersen SE (1996): PET activation of posterior temporal regions during auditory word presentation and verb generation. Cereb Cortex 6:1–10.

Friederici AD, Meyer M, von Cramon DY (2000a): Auditory language comprehension: an event-related fMRI study on the processing of syntactic and lexical information. Brain Lang 75:465–477.

Friederici AD, Wang Y, Herrmann CS, Maess B, Oertel U (2000b): Localization of early syntactic processes in frontal and temporal cortical areas: a magnetoencephalographic study. Hum Brain Mapp 11:1–11.

Friston KJ (1994): Statistical parametric maps in functional imaging: a general linear approach. Hum Brain Mapp 2:189–210.

Gandour J, Wong D, Hsieh L, Weinzapfel B, Van Lancker D, Hutchins GD (2000): A cross linguistic PET study of tone perception. J Cogn Neurosci 12:207–222.

Giraud AL, Kell C, Klinke R, Russ MO, Sterzer P, Thierfelder C, Preibisch C, Kleinschmidt A (2001): Dissociating effort and success in speech comprehension: an fMRI study. Neuroimage 13: S534.

Goodglass H (1993): Understanding aphasia. San Diego: Academic Press.

Heilman KM, Bowers D, Speedie L, Coslett HB (1984): Comprehension of affective and non affective prosody. Neurology 84:917–921.

Hickok G, Pöppel D (2000): Towards a functional neuroanatomy of speech perception. Trends Cogn Sci 4:131–138.

Humphries C, Buchsbaum C, Hickok G (2001): Role of anterior temporal cortex in auditory sentence comprehension: an fMRI study. Neuroreport 12:1749–1752.

Huynh H, Feldt LS (1970): Conditions under which the mean square ratios in repeated measurements designs have exact F distributions. Journal of the American Statistical Association 65:1582–1589.

Jäncke L, Shah NJ, Posse S, Grosse-Ruyken M, Müller-Gärtner H (1998): Intensity coding of auditory stimuli: an fMRI study. Neuropsychologia 36:875–883.

Johnsrude IS, Penhune B, Zatorre RJ (2000): Functional specificity in the right human auditory cortex for perceiving pitch direction. Brain 123:155–163.

Kuperberg GR, McGuire PK, Bullmore ET, Brammer MJ, Rabe-Hesketh S, Wright IC, Lythgoe DJ, Williams SCR, David AS (2000): Common and distinct neural substrates for pragmatic, semantic, and syntactic processing of spoken sentences: an fMRI study. J Cogn Neurosci 12:321–341.

Lee JH, Garwood M, Menon R, Adriany G, Andersen P, Truwit CL, Ugurbil K (1995): High contrast and fast three-dimensional magnetic resonance imaging at high fields. Magn Reson Med 34:308.

Lohmann G, Müller K, Bosch V, Mentzel H, Hessler S, Chen L, Zysset S, von Cramon DY (2001): Lipsia—a new software system for the evaluation of functional magnetic resonance images of the human brain. Comput Med Imaging Graph 25:449–457.

Mazoyer BM, Tzourio N, Frak V, Syrota A, Murayama N, Levrier O, Salamon G, Dehaene S, Cohen L, Mehler J (1993): The cortical representation of speech. J Cogn Neurosci 5:467–479.

Meyer M, Friederici AD, von Cramon DY (2000): Neurocognition of auditory sentence comprehension: event-related fMRI reveals sensitivity to syntactic violations and task demands. Brain Res Cogn Brain Res 9:19–33.

Müller RA, Rothermel RD, Behen ME, Muzik O, Mangner TJ, Chugani HT (1997): Receptive and expressive language activations for sentences: a PET study. Neuroreport 8:3767–3770.

Norris DG (2000): Reduced power multi-slice MDEFT imaging. Magn Reson Imaging 11:445–451.

Paulesu E, Frith CD, Frackowiak RSJ (1993): The neural correlates of the verbal components of working memory. Nature 362:342–345.

Pell MD (1998): Recognition of prosody following unilateral brain lesion: influence of functional and structural attributes of prosodic contours. Neuropsychologia 36:701–715.

Pell MD, Baum SR (1997): The ability to perceive and comprehend intonation in linguistic and affective contexts by brain-damaged adults. Brain Lang 57:80–89.

Perkins JM, Baran JA, Gandour J (1996): Hemispheric specialization in processing intonation contours. Aphasiology 10:343–362.

Perry DW, Zatorre RJ, Petrides M, Alivisatos B, Meyer E, Evans AC (1999): Localization of cerebral activity during simple singing. Neuroreport 11:3979–3984.

Poldrack RA, Temple E, Protopapas A, Nagarajan S, Tallal P, Merzenich M, Gabrieli JDE (2001): Relations between the neural bases of dynamic auditory processing and phonological processing: evidence from fMRI. J Cogn Neurosci 13:687–697.

Poldrack RA, Wagner AD, Prull MW, Desmond JE, Glover GH, Gabrieli JDE (1999): Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. Neuroimage 10:15–35.

Pöppel D (1996): A critical review of PET studies of phonological processing. Brain Lang 55:317–351.

Price PJ, Ostendorf M, Shattuck-Hufnagel S, Fong C (1991): The use of prosody in syntactic disambiguation. J Acoust Soc Am 90:2956–2970.

Riecker A, Ackermann H, Wildgruber D, Dogil G, Grodd W (2000): Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. Neuroreport 11:1997–2000.

Ross ED, Thompson RD, Yenkosky J (1997): Lateralization of affective prosody in brain and the callosal integration of hemispheric language functions. Brain Lang 56:27–54.

Schirmer A, Alter K, Kotz SA, Friederici AD (2001): Lateralization of prosody during language production: a lesion study. Brain Lang 76:1–17.

Schlosser MJ, Aoyagi N, Fulbright RK, Gore JC, McCarthy G (1998): Functional MRI studies of auditory comprehension. Hum Brain Mapp 6:1–13.

Scott SK, Blank CC, Rosen S, Wise RJS (2000): Identification of a pathway for intelligible speech in the left temporal lobe. Brain 123:2400–2406.

Shattuck-Hufnagel S, Turk AE (1996): A prosody tutorial for investigators of auditory sentence processing. J Psycholinguist Res 25:193–247.

Sonntag GP, Portele T (1998): PURR—a method for prosody evaluation and investigation. J Comput Speech Lang 12:437–451.

Starkstein SE, Federoff JP, Price TR, Leiguarda RC, Robinson RG (1994): Neuropsychological and neuroradiological correlates of emotional prosody comprehension. Neurology 44:515–522.

Steinhauer K, Alter K, Friederici AD (1999): Brain potentials indicate immediate use of prosodic cues in natural speech processing. Nat Neurosci 2:191–196.

Steinhauer K, Friederici AD (2001): Prosodic boundaries, comma rules, and brain responses: the closure positive shift in ERPs as a universal marker for prosodic phrasing in listeners and readers. J Psycholinguist Res 30:267–295.

Stowe L, Cees A, Broere A, Paans A, Wijers A, Mulder G, Vaalburg W, Zwarts F (1998): Localizing components of a complex task: sentence processing and working memory. Neuroreport 9:2995–2999.

Talairach J, Tournoux P (1988): Co-planar stereotaxic atlas of the human brain. New York: Thieme.

Thirion JP (1998): Image matching as a diffusion process: an analogy with Maxwell's daemons. Med Image Anal 2:243–260.

Tzourio N, Massioui FE, Crivello F, Joliot M, Renault B, Mazoyer B (1997): Functional anatomy of human auditory attention studied with PET. Neuroimage 5:63–77.

Ugurbil K, Garwood M, Ellermann J, Hendrich K, Hinke R, Hu X, Kim SG, Menon R, Merkle H, Ogawa S, et al. (1993): Imaging at high magnetic fields: initial experiences at 4 T. Magn Res Q 9:259–277.

van Lancker D, Sidtis JJ (1992): The identification of affective-prosodic stimuli by left- and right-hemisphere-damaged subjects: all errors are not created equal. J Speech Hear Res 35:963–970.

Weintraub S, Mesulam MM (1981): Disturbances of prosody. A right-hemisphere contribution to language. Arch Neurol 38:742–744.

Wessinger CM, Van Meter J, Tian B, Pekar J, Rauschecker JP (2001): Hierarchical organization of the human auditory cortex revealed by functional magnetic resonance imaging. J Cog Neurosci 13:1–7.

Wise RJS, Chollet F, Hadar U, Friston KJ, Hoffner E, Frackowiak RSJ (1991): Distribution of cortical neural networks involved in word comprehension and word retrieval. Brain 114:1803–1817.

Zarahn E, Aguirre G, D'Esposito M (1997): Empirical analysis of BOLD-fMRI statistics. I. Spatially smoothed data collected under null-hypothesis and experimental conditions. Neuroimage 5:179–197.

Zatorre RJ, Belin P (2001): Spectral and temporal processing in human auditory cortex. Cereb Cortex 11:946–953.

Zatorre RJ, Evans AC, Meyer E (1994): Neural mechanisms underlying melodic perception and memory for pitch. J Neurosci 14:1908–1919.

Zatorre RJ, Evans AC, Meyer E, Gjedde A (1992): Lateralization of phonetic and pitch discrimination in speech processing. Science 256:846–849.

Zatorre RJ, Samson S (1991): Role of the right temporal neocortex in retention of pitch in auditory short-term memory. Brain 114:2403–2417.

Zurif EB (1974): Auditory lateralization: prosodic and syntactic factors. Brain Lang 1:391–404.