

On the lateralization of emotional prosody: An event-related functional MR investigation

Sonja A. Kotz,^{a,*} Martin Meyer,^b Kai Alter,^a Mireille Besson,^c D. Yves von Cramon,^a and Angela D. Friederici^a

^a *Max-Planck-Institute of Cognitive Neuroscience, Stephanstrasse 1a, P.O. Box 500355, Leipzig D-04317, Germany*

^b *Center for Functional Imaging Studies, University of Edinburgh, UK*

^c *Centre National de la Recherche Scientifique (CNRS), Marseille, France*

Accepted 23 July 2002

Abstract

In order to investigate the lateralization of emotional speech we recorded the brain responses to three emotional intonations in two conditions, i.e., “normal” speech and “prosodic” speech (i.e., speech with no linguistic meaning, but retaining the ‘slow prosodic modulations’ of speech). Participants listened to semantically neutral sentences spoken with a positive, neutral, or negative intonation in both conditions and judged how positive, negative, or neutral the intonation was on a five-point scale. Core peri-sylvian language areas, as well as some frontal and subcortical areas were activated bilaterally in the normal speech condition. In contrast, a bilateral fronto-opercular region was active when participants listened to prosodic speech. Positive and negative intonations elicited a bilateral fronto-temporal and subcortical pattern in the normal speech condition, and more frontal activation in the prosodic speech condition. The current results call into question an exclusive right hemisphere lateralization of emotional prosody and expand patient data on the functional role of the basal ganglia during the perception of emotional prosody.

© 2003 Elsevier Science (USA). All rights reserved.

Keywords: fMRI; Emotional prosody; Lateralization; Linguistic prosodic cues vs. non-linguistic prosodic cues

1. Introduction

A persistent question in the investigation of auditory language processing is the influence of linguistic and non-linguistic (emotional) prosodic cues on language comprehension. Linguistic prosodic cues, for example, are used during sentence processing to establish a syntactic structure in a perceived speech signal. The evidence for this point is not extensive but is convincing. Behavioral and event-related brain potential studies (ERPs) in healthy populations suggest an interaction of linguistic prosodic cues with language-specific subprocesses, such as syntax (e.g., Cutler, Dahan, & van Donselaar, 1997; Steinhauer, Alter, & Friederici, 1999; Warren, Grabe, & Nolan, 1995). Regarding the role of non-linguistic prosodic cues that correlate with emotional prosody in language comprehension, results of

recent brain potential studies (Kotz, Alter, Besson, Schirmer, & Friederici, 2000; Pihan, Ackermann, & Altenmüller, 1997, 2000) and of two brain imaging studies (Buchanan et al., 2000; George et al., 1996). All of these studies above indicate that emotional prosody may be processed in the right hemisphere (Buchanan et al., 2000; George et al., 1996), but may also induce bilateral activation patterns as a function of subvocal rehearsal (Pihan et al., 2000) or task type (Kotz et al., 2000).

The lateralization of linguistic or emotional speech prosody has been extensively investigated in patients, but the results are mixed. While some data indicate that both emotional and linguistic prosody are processed in the right hemisphere (e.g., Bryan, 1989; Dykstra, Gandour, & Stark, 1995), other results suggest that only emotional prosody (e.g., Blonder, Bowers, & Heilman, 1991; Borod, 1993; Starkstein, Federoff, Price, Leiguarda, & Robinson, 1994) or only linguistic prosody (e.g., Brådvik et al., 1991; Weintraub, Mesulam, &

* Corresponding author. Fax: +49-3-41-99-40-113.

E-mail address: kotz@cns.mpg.de (S.A. Kotz).

Kramer, 1981) is processed in the right hemisphere. Note also that some results imply that sentence-level linguistic prosody is processed in the left hemisphere (e.g., Emmorey, 1987; Van Lancker, 1980).

There has also been disagreement whether emotional valence induces lateralization of emotional prosody. The lateralization hypothesis implies a right hemisphere superiority of emotional speech independent of valence (e.g., Ross, 1981; Ross, Thompson, & Yenkosky, 1997). By contrast, the valence hypothesis proposes a left hemisphere dominance for positive states and a right hemisphere dominance for negative states (e.g., Davidson, Abercrombie, Nitschke, & Putnam, 1999). However, based on their finding of no consistent lateralization effects of emotional valence, Pell and Baum (1997) call into question the emotional valence hypothesis.

Alternatively, it has been proposed that the lateralization of prosodic processing may vary as a function of the acoustic parameters of prosody, such as fundamental frequency (F_0), intensity or duration, under study. For example, it has been suggested that F_0 is processed in the right hemisphere, while intensity and duration are processed in the left hemisphere (e.g., Van Lancker & Sidtis, 1992; Zatorre, Belin, & Penhune, 2002). Note moreover, that these lateralization patterns are possibly independent of whether prosody is emotional or linguistic (e.g., Ouellette & Baum, 1993; Zatorre, 1988).

As this brief review of the literature shows, neither patient studies nor brain imaging studies have yet provided a clear picture of the lateralization of linguistic and emotional prosody (see also Baum & Pell, 1999; Pell, 1998; Ross et al., 1997). In particular, two issues have been critically discussed. First, Pell (1998) has argued that in most patient studies, the interactive influence of linguistic and emotional factors on speech prosody has not been controlled for. Second, Ross et al. (1997) have pointed out that patients with different sizes and areas of lesions have been included in the studies on aprosodia, thus leading to heterogeneous results and complicating any meaningful conclusion on the lateralization of emotional prosody.

The overall aim of the present study is to shed light on the lateralization of emotional prosody by using a new brain imaging method, event-related fMRI. Specifically, we focussed on two aspects. First, we compared normal and prosodic speech. In the latter condition, the speech signal was filtered so as to suppress segmental and lexical information but to keep suprasegmental information intact. Thus, while the acoustic parameters of emotional prosody are kept constant across normal and prosodic speech, no lexical, semantic or syntactic information is available in prosodic speech. It is as if one is listening to a human voice speaking behind closed doors. Such a comparison therefore allows to investigate which hemisphere and brain areas are more sensitive to: (a) language-specific cues, such as lexical-semantic and

syntactic aspects in normal speech and (b) purely prosodic cues as in prosodic speech. Based on previous evidence (e.g., Meyer, Alter, Friederici, Lohmann, & von Cramon, 2002), we predicted bilateral temporal activation and frontal activation in both conditions, but we expected the frontal activation to be weaker in the normal speech condition. In the prosodic speech condition we expected weaker temporal activation, but enhanced frontal activation. In addition, the prosodic speech condition should reveal which hemisphere responds more sensitive to slow prosodic modulations of speech. Second, while the semantic content of the sentences was always neutral, the sentences were spoken with positive, negative or neutral intonations. It was of interest to determine whether different brain areas will be activated by these different types of intonations and whether results would be the same or different for normal and prosodic speech.

2. Methods

2.1. Subjects

Twelve right-handed native German speakers (eight female) with a mean age of 24 years (range: 22–29 years) with normal hearing skills participated in the experiment upon written consent according to the guidelines of the Ethics Committee of the University of Leipzig Medical Faculty.

2.2. Stimuli

Stimuli consisted of a total of 108 sentences that were spoken by a trained female speaker of German in a sound proof room (IAC) at a 16 bit/44.1 kHz sampling rate. Sentences were then digitized, downsampled at a 16 bit/16 kHz sampling rate and normalized in amplitude (70%). All sentences had similar syntactic structure and length, they all started with personal pronouns and their semantic content was always neutral (i.e., *Sie hat die Zeitung gelesen* [non-literal translation: *She has read the newspaper*]). A third of the sentences (36) were spoken with positive intonation (i.e., a happy voice), a third with negative intonation (i.e., an angry voice), and a third with neutral intonation. In the prosodic speech condition, these 108 sentences were delexicalized by applying the PURR-filtering procedure (Sonntag & Portele, 1998). By filtering out all acoustic information above the third harmonic and all aperiodic signals (see Figs. 1 and 2), this procedure leaves suprasegmental information, such as the global F_0 contour and amplitude intact, but suppresses all the segmental and lexical information. Note that this procedure is more efficient in suppressing phonological and lexical information than low pass filters. Thus, the listener only perceives slow

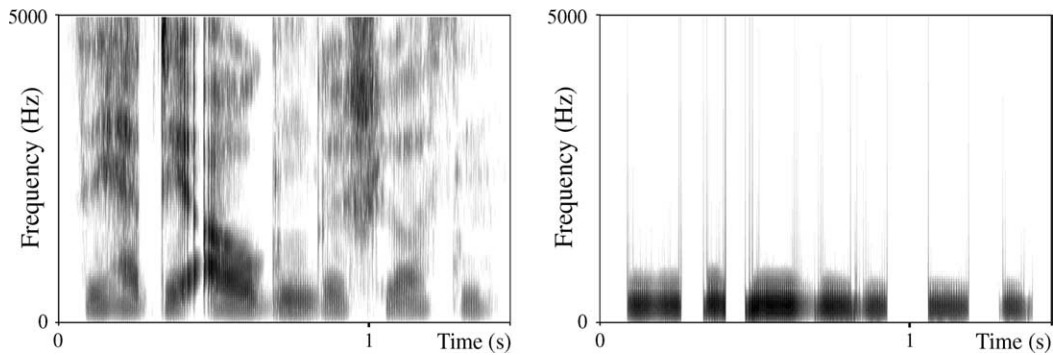


Fig. 1. Wide band spectrogram of speech signals before and after application of the PURR-filter. The left spectrogram illustrates the frequency spectrum (0–5 kHz) of a normal sentence. The right image illustrates clearly the reduced spectral information derived from a PURR-filter treated prosodic speech stimulus. The acoustic signal derived from this filtering procedure is reduced to frequencies containing the F_0 as well as second and third harmonic. Additionally, all aperiodic portions of the speech signal are removed from the speech signal.

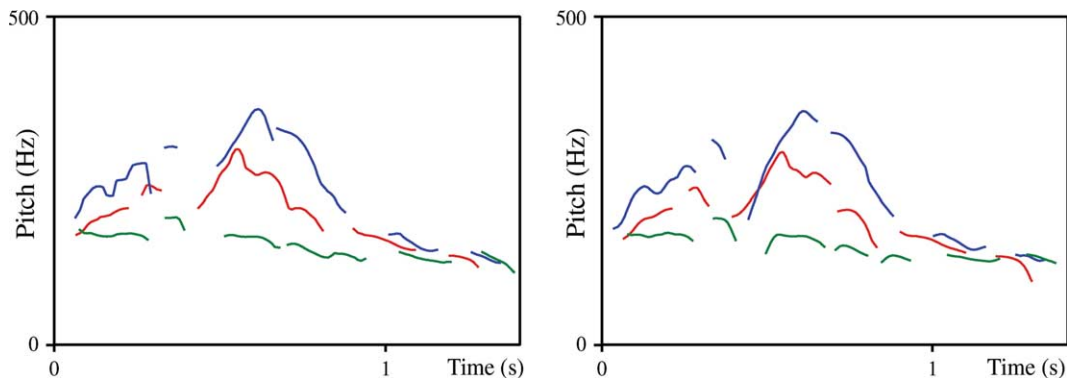


Fig. 2. Pitch contour of speech signals before and after application of the PURR-filter. The left image illustrates the pitch contours of a sentence in the normal speech condition. The blue line deflects the positive contour, the red line the negative contour, and the green line the neutral contour. Peaks and valleys of pitch contour symbolize typical sentence intonation. The right image shows pitch contours (positive, negative, and neutral) for the same sentence in the prosodic speech condition suggesting that the filtering procedure does not change the intonation contour.

prosodic modulations as pitch variations over time, but no intelligible linguistic information. These stimuli are called prosodic speech. To summarize, the stimulus materials comprised a total of 216 sentences in two conditions of speech prosody (normal speech and prosodic speech) with three levels of emotional intonation (positive, negative, and neutral).

2.3. Procedure

Before the experimental session in the scanner, participants were instructed and given a brief training with feedback to ensure the judgment of emotional intonations in both the normal and the prosodic speech conditions. In the scanner participants were presented with 216 sentences (36 in each of the six conditions) in a pseudo-randomized order. Sentences were not repeated during the experiment. The sounds were presented binaurally via headphones that were specifically adapted for use in an fMRI environment. A combination of external ear defenders and perforated ear plugs that conducted the

sound directly into the auditory channel, was used to attenuate the scanner noise without reducing the quality of speech stimulation. Participants were asked to judge via key presses immediately after the presentation of each normal or prosodic sentence whether it was spoken with neutral, positive, or negative intonation. In the scanner no feedback was given. The design assigned a fixed presentation rate and each trial started with the presentation of one single sentence. The average sentence length was 1500 ms. The successive presentation of single sentences was separated by an inter-trial-interval of 7500 ms.

2.4. 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. The experiment consisted of two separate, but consecutive sessions. In the first session, high resolution whole-head 3D MDEFT brain scans (128 sagittal slices, 1.5 mm thickness, FOV $25.0 \times 25.0 \times 19.2$ cm, data

matrix of 256×256 voxels) were acquired for reasons of improved localization (Lee et al., 1995; Ugurbil et al., 1993). The second session started with the collection of scout spin echo sagittal scans to define the anterior and posterior commissures on a midline sagittal section. For each participant, 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 were extended dorsally. The whole range of slices comprised an anatomical volume of 46 mm and covered all parts of the peri-sylvian cortex and extended dorsally to the intraparietal sulcus. For functional imaging, a gradient-echo EPI sequence was used with a TE of 30 ms, a flip angle of 90° , a TR of 2000 ms, and an acquisition bandwidth of 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 and to reduce noise while participants listen to the sentences. The matrix acquired was 64×64 with a FOV of 19.2 cm, resulting in an in-plane resolution of 3×3 mm. The slice thickness was 5 mm with an interslice gap of 2 mm. To align the functional EPI images to 3D-MDEFT images, conventional T_1 weighted, MDEFT, and T_1 weighted EPI images were obtained in-plane with the T_2^* echo-planar images as reference.

2.5. Data processing

The data processing was performed using the software package LIPSIA (Lohmann et al., 2001). Functional data were corrected for slicetime acquisition differences using sinc-interpolation. In addition, the data were corrected for motion artifacts. A temporal high-pass filter with a cutoff frequency of 1/36 Hz was used for baseline correction of the signal, and a spatial gaussian filter with $\sigma = 0.8$ was applied. The increased autocorrelation due to filtering was taken into account during statistical evaluation.

To align the functional dataslices with a 3D stereotactic coordinate reference system, a rigid linear registration with six degrees of freedom (3 rotational, 3 translational) was performed. The rotational and translational parameters were acquired on the basis of the MDEFT and EPI- T_1 slices to achieve an optimal match between these slices and the individual 3D reference data set. The MDEFT volume data set with 160 slices and 1 mm slice thickness was standardized to the Talairach stereotactic space (Talairach & Tournoux, 1988). The rotational and translational parameters were subsequently transformed by linear scaling to a standard size. The resulting parameters were then used to transform the functional slices using trilinear interpolation so that the resulting functional slices were aligned with the stereotactic coordinate system.

The statistical evaluation was based on a least-squares estimation using the general linear model for serially autocorrelated observations (Bosch, 2000; Friston, 1994a,b; Friston et al., 1995a; Friston et al., 1995b; Worsley & Friston, 1995; Zarahn, Aguirre, & D'Esposito, 1997). The design matrix was generated with a synthetic hemodynamic response function and a response delay of 6 s (Friston et al., 1998; Josephs, Turner, & Friston, 1997). The model equation, including the observation data, the design matrix, and the error term, was convolved with a Gaussian kernel of dispersion of 4 s.

Five contrasts were calculated for each participant. The GLM allows for the calculation of collapsed contrasts, i.e., the generation of statistical parametric maps (SPMs) comparing global brain activation during the perception of normal speech vs. prosodic speech irrespective of the particular emotional intonation. Moreover, for each participant, the contrasts between different intonations (i.e., neutral vs. positive, neutral vs. negative, in both normal and prosodic speech) were calculated using the t -statistics. Subsequently, t -values were converted to Z scores. To achieve a better signal-to-noise ratio, intersubject averaging of individual SPMs was done for each contrast separately.

Local maxima of the SPMs are listed in Tables 2–4. A pixel was defined to be a local maximum if its z -value exceeded $|3.09|$ and if it was the largest within a 5 mm radius. Local maxima residing in activation areas smaller than 100 mm^3 are not reported.

3. Results

3.1. Behavioral data

As can be seen in Table 1, participants responded to the intonation of the sentences with a high accuracy rate in all three normal speech conditions. In contrast, the responses in the prosodic speech conditions were lower than in normal speech for the positive and negative intonations. For the neutral intonation, results were similar in both the normal and prosodic speech conditions.

3.2. Imaging data

Five contrasts were calculated to reveal common or differential activation patterns between (i) normal and prosodic speech, irrespective of intonation, (ii) neutral vs. positive and neutral vs. negative intonations in normal speech, and (iii) neutral vs. positive and neutral vs. negative intonations in prosodic speech.

Normal vs. prosodic speech. Comparing normal and prosodic speech across all three emotional intonations revealed bilateral, but left accentuated, activation of temporal and subcortical regions (putamen and thalamus)

Table 1
Behavioral data

Prosody	Normal speech (%)	Standard error (\pm)	Prosodic speech (%)	Standard error (\pm)
Positive	99	2.55	88	1.76
Neutral	98	2.70	98	2.51
Negative	99	1.47	84	1.84

Prosodic judgment performance for each of the two prosodic speech conditions and the different emotional contours. Data are presented as % correct and standard error.

as well as activations of left inferior frontal regions for normal speech. Prosodic speech resulted in bilateral inferior frontal, prefrontal, and subcortical (caudate) activation (see Fig. 3, Table 2).

Prosodic effects in normal speech. Both the neutral vs. positive and the neutral vs. negative contrasts elicited

bilateral activation in frontal, temporal, and subcortical (caudate) regions. Overall, the activation was stronger for positive intonations than negative intonations. However, negative intonations were associated with stronger activation of the bilateral rolandic operculum/insula regions (see Fig. 4, Table 3).

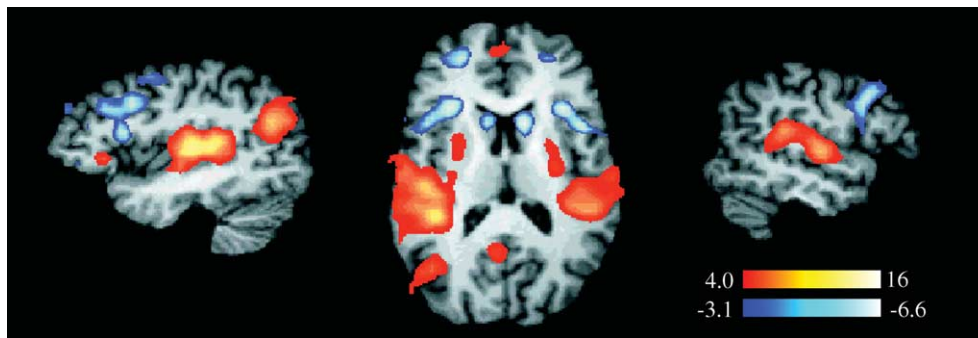


Fig. 3. Normal speech vs. prosodic speech. Functional intersubject activation ($N = 12$) exceeding the significant threshold is shown in a left sagittal, axial, and right sagittal view. The contrast between normal speech (red) and prosodic speech (blue) across the three prosodic contours revealed bilateral activation of temporal and subcortical regions for the normal speech condition and bilateral prefrontal and inferior frontal activation for the prosodic speech condition. Functional activation was thresholded at $Z \geq 4.0$ for normal speech and $Z \leq -3.1$ for prosodic speech.

Table 2
Normal speech vs. prosodic speech

Location	Z score	Left hemisphere			Z score	Right hemisphere		
		x	y	z		x	y	z
Normal speech > prosodic speech								
Fronto-median cortex	7.28	-4	58	13	-	-	-	-
IFG	6.72	-43	23	-6	-	-	-	-
STR	16.84	-52	-17	6	10.40	55	-12	2
AG	9.62	-41	-60	32	-	-	-	-
CG	6.69	-5	-51	33	-	-	-	-
Putamen	6.45	-26	3	1	8.30	22	0	-2
Thalamus	4.73	-11	-30	5	3.90	34	-58	43
Prosodic speech > normal speech								
Fronto-striatal cortex	-6.93	-23	48	-5	-4.87	25	48	-5
IFS/MFG	-5.20	-47	18	23	-5.40	37	5	23
Fronto-opercular cortex	-6.60	-32	-22	2	-5.90	31	23	5
Caudate (Head)	-4.92	-11	13	5	-6.14	10	14	8

Tables 2–4, respectively list the results of direct comparisons between conditions. Z scores indicate the magnitude of statistical significance. Localization is based on stereotactic coordinates (Talairach & 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. Thresholds for functional activation were set at $|Z| \geq 3.1$. The table only lists activation clusters exceeding a minimal size of 150 voxels. Anatomical locations are abbreviated as follows: IFG, inferior frontal gyrus, MFG, middle frontal gyrus, SFG, superior frontal gyrus, aINS, anterior insula, INS, insula, ROP, rolandic operculum, HG, Heschl's gyrus, STS, superior temporal sulcus, MTG, middle temporal gyrus, PT, planum temporale, IPL, inferior parietal lobe, IPS, intraparietal sulcus, STR, superior temporal region, AG, angular gyrus, CG, cingulate gyrus, IFS, inferior frontal sulcus.

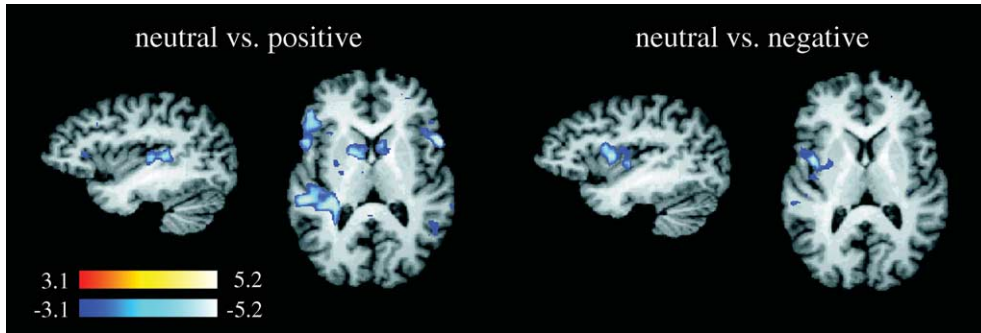


Fig. 4. Prosodic effects–normal speech. The activation patterns for the positive effect (left) and the negative effect (right) are displayed in a left sagittal and an axial view. No significant activation could be detected for neutral contour (red), but several brain regions were more strongly involved in processing positive/negative contours (blue). Functional activation was thresholded at $Z \geq 3.1$ for neutral and $Z \leq -3.1$ for positive/negative contours.

Table 3
Prosodic effects–normal speech

Location (BA)	Z score	Left hemisphere			Z score	Right hemisphere		
		x	y	z		x	y	z
Neutral < positive								
IFG (44)	-4.2	-50	13	1	-5.2	52	13	4
IFG (45/47)	-4.4	-47	28	1	-	-	-	-
Medial HG	-4.3	-44	-22	9	-	-	-	-
PT	-4.8	-53	-30	11	-	-	-	-
Post. STS/MTG	-4.5	-44	-49	9	-4.1	46	-52	12
IPL (40)	-	-	-	-	-3.7	46	-46	26
Caudate	-4.0	-7	2	4	-3.7	5	8	2
Neutral < negative								
ROP/INS	-4.3	-41	3	8	-3.8	32	-5	16
Lateral HG	-3.5	-58	-13	12	-	-	-	-

Functional activation indicated separately for contrasts between conditions. For explanations see Table 2.

Prosodic effects in prosodic speech. Both the neutral vs. positive and the neutral vs. negative contrasts revealed similar activation of frontal, temporal, and subcortical regions. While both the frontal IFG and subcortical activations were bilateral, the overall pattern of activation was more left than right lateralized for both contrasts (see Fig. 5, Table 4).

4. Discussion

The current experiment set out to investigate whether a controlled manipulation of linguistic information would help to specify the underlying neural mechanisms of emotional prosody perception. We attempted to define to which extent emotional prosody is only lateralized

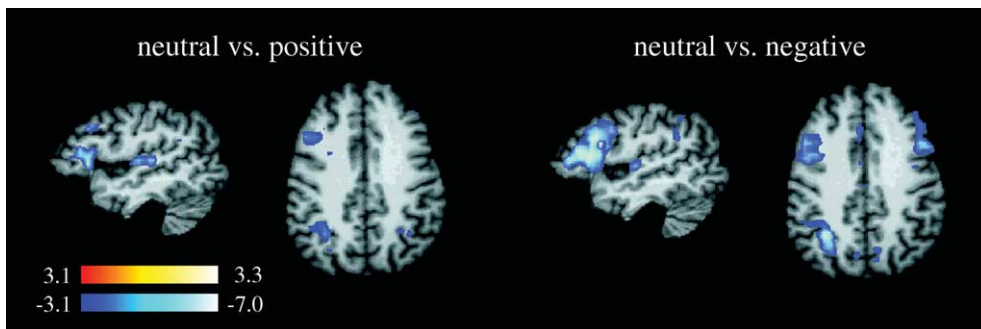


Fig. 5. Prosodic effects–prosodic speech. The activation patterns for the positive effect (left) and the negative effect (right) are displayed in a left sagittal and an axial view. No significant activation could be detected for neutral contour (red), but several brain regions were more strongly involved in processing positive/negative contours (blue). Functional activation was thresholded at $Z \geq 3.1$ for neutral and $Z \leq -3.1$ for positive/negative contours.

Table 4
Prosodic effects–prosodic speech

Location (BA)	Z score	Left hemisphere			Z score	Right hemisphere		
		x	y	z		x	y	z
Neutral < positive								
aINS	–3.3	–28	22	0	–	–	–	–
IFG (44/45)	–5.2	–53	19	4	–3.5	44	21	13
IFG (47)	–3.5	–47	35	–2	–	–	–	–
MFG (46)	–	–	–	–	–4.3	28	38	6
Lateral HG	–4.8	–49	–14	7	–	–	–	–
PT	–3.9	–55	–34	12	–	–	–	–
IPS	–5.3	–32	–46	34	–	–	–	–
Caudate	–3.9	–9	19	3	–4.2	10	16	1
Neutral < negative								
aINS	–5.4	–29	22	0	–	–	–	–
IFG (44/45)	–6.3	–44	21	13	–	–	–	–
IFG (45)	–5.4	–52	21	3	–	–	–	–
IFG (47)	–5.1	–47	35	–2	–	–	–	–
IFG/MFG (45/46)	–5.8	–44	32	7	–4.4	40	32	11
SFG (10)	–	–	–	–	–4.0	40	19	20
Lateral HG	–4.2	–48	–14	7	–	–	–	–
PT	–4.2	–54	–32	11	–	–	–	–
IPS	–5.8	–31	–48	40	–	–	–	–
Caudate	–5.3	–9	11	5	–5.5	11	15	1

Functional activation indicated separately for contrasts between conditions. For explanations see Table 2.

in the right hemisphere and explored whether emotional intonations (i.e., positive, negative, and neutral) in prosodic speech can be differentiated. The present data provide evidence for differential brain activation patterns that vary as a function of normal and prosodic speech. In addition, mainly overlapping brain activation was found when listening to positive and negative emotional intonations in normal and prosodic speech with slight, but noteworthy differences. In the following discussion, critical contrasts will be discussed in turn focussing on the main activation areas, namely frontal, subcortical and temporal activation in the context of lateralization.

4.1. Frontal activation

Normal vs. prosodic speech. The activation of the frontal cortex and any related lateralization is currently a matter of debate. Studies reporting frontal activation for normal speech argue that the anterior part of the left inferior frontal gyrus (BA 45) in particular reflects lexical semantic processes (e.g., Poldrack et al., 1999), while others imply that the dorsal and posterior part of the left IFG organizes syntactic processing (Caplan, Alpert, & Waters, 1999; Carpentier et al., 2001; Dapretto & Bookheimer, 1999; Indefrey, Haggort, Herzog, Seitz, & Brown, 2001). However, given that in the present study we find frontal activation in both normal speech (left-lateralized) and prosodic speech (bilateral, but left-accentuated), one cannot conclude that the frontal activation pattern solely reflects the processing of specific linguistic information

(syntax or lexico-semantic), as this information was eliminated in the prosodic speech condition. This view is supported by a recent study demonstrating the involvement of the left inferior frontal cortex during the processing of music (Maess, Koelsch, Gunter, & Friederici, 2001).

An alternative hypothesis is that the bilateral, but left-accentuated frontal activation in prosodic speech reflects effortful processing of prosodic information. In the current experiment, the discrimination of emotional intonations correlated with increased task demands in prosodic speech. Second, due to the event-related design trials of both, the normal and the prosodic speech conditions were presented in a pseudo-randomized manner. Thus, switching between the two conditions could have enhanced cognitive demands causing an increase of the frontal activation in both conditions. Enhanced activation of the left IFG with increasing computational demands in a phoneme monitoring task (Demonet, Price, Wise, & Frackowiak, 1994) and in a phonological discrimination task (Burton, Small, & Blumstein, 2000) was previously shown. Therefore, the left accentuated frontal activation pattern in prosodic speech might reflect the effort to discriminate phonological aspects in different auditory modes, as the main difference between normal and prosodic speech is speech intelligibility. As the relative importance of prosodic parameters, such as pitch and amplitude is likely to be enhanced in the absence of lexical information, a bilateral fronto-opercular activation might reflect processing associated with the extraction of prosodic cues under

increased task demands such as the mixed presentation of intelligible and unintelligible speech. Similar evidence was recently reported for a comparison of incomprehensible and comprehensible speech (Giraud et al., 2001; Poldrack et al., 2001).

Prosodic effects in normal and prosodic speech. The activation of the IFG for positive intonations in normal speech and both positive and negative intonations in prosodic speech overlap with the frontal activation reported for the overall contrast of normal and prosodic speech. These data suggest that the discrimination of emotional intonations, mainly independent of emotional valence, does result in bilateral activation of the frontal cortex, in particular Brodmann's area 44, which has been linked to phonological discrimination (Burton et al., 2000, but left lateralized). Thus, a combination of task (discrimination of prosodic intonations) and task demands (normal vs. prosodic speech) could have induced the frontal activation pattern in the current fMRI experiment. However, the hemodynamic changes elicited by negative emotional intonations in normal speech show a slightly different picture. Here comparable to positive and negative intonations in prosodic speech, left-accented activation of the insula and Heschl's gyrus was found, but in addition bilateral activation of the rolandic operculum. The role of the left and right insula in speech prosody has been investigated in patient studies on aprosodic syndromes (Cancelliere & Kertesz, 1990) though further supportive evidence is rather sparse. The current data do support a role for the insula in the perception of emotional prosody. An interpretation of the activation in the rolandic operculum, however, must remain speculative as there is no evidence on its role in the perception of speech prosody. For speech production, however, there is fMRI data in which the rolandic operculum shows activation when comparing overt and covert speech (Riecker, Ackermann, Wildgruber, Dogil, & Grodd, 2000). A link between the rolandic operculum activation in these two studies can only be made due to two assumptions: (a) inner speech is only used for the discrimination of negative, but not for positive intonation, as it is more difficult (b) inner speech is more likely to be applied when the perceptual input contains words, as in the normal speech condition.

4.2. Subcortical activation

Normal vs. prosodic speech. Specific areas of the subcortical network were activated in both, but more extended in normal than prosodic speech, involving bilateral activation of the putamen, thalamus and the head of the caudate. Recently, Hall et al. (2000) speculated that the thalamus serves as a gating mechanism for the transmission of sensory information to the cortex and is involved in auditory attention by tuning sensory neurons in the auditory cortex. Furthermore, Cancelliere

and Kertesz (1990) reported that patients with right and left hemisphere lesions of the striatal structure (caudate nucleus and putamen) display aprosodic syndromes. In addition, patients with Huntington's chorea which results from cell loss in the striatum (Cancelliere & Hausdorf, 1987; Ross, D., Edmondson, Seibert, & Homan, 1988) and patients with Parkinson's disease (Blonder, Gur, & Gur, 1989; Scott, Caird, & Williams, 1984) which results from a degeneration of dopaminergic innervations in the striatal complex (Nauta, 1986) can also suffer from aprosodia. Therefore, the current data support previously reported patient data that imply a clearly bilateral functional role of parts of the basal ganglia in the perception of emotional prosody (see Ross, 1981 for alternative results).

Prosodic effects in normal and prosodic speech. We report bilateral, but partial striatal (caudate) activation comparing positive and negative intonations in the normal and prosodic speech condition (for a possible explanation of negative intonations in normal speech see discussion on frontal activation above). Once again, discrimination of emotional prosody engages parts of the basal ganglia and this is mainly independent of emotional valence (see also Cancelliere & Kertesz, 1990).

4.3. Temporal activation

Normal vs. prosodic speech. The temporal activation (superior temporal region; STR) is in line with previous fMRI results on spoken words (e.g., Binder et al., 2000) and sentence comprehension (e.g., Dehaene et al., 1997; Friederici, Meyer, & von Cramon, 2000; Meyer et al., 2002; Meyer, Friederici, & von Cramon, 2000) as well as studies on the perceptual analysis of speech signals (Hickok & Poeppel, 2000; Zatorre, Evans, Meyer, & Gjedde, 1992). Activation of the STR was extremely reduced for prosodic speech. This result is in accordance with the prediction that when the speech signal is stripped of phonological and lexical-semantic information and is linguistically unintelligible, there should be a decrease of activation in auditory cortices (Poldrack et al., 2001). However, a clear right STR lateralization for prosodic vs. normal speech is not supported by the current results (e.g., Meyer, Alter, Steinhauer, & Friederici, 2001).

Prosodic effects in normal speech and prosodic speech. In examining the specific emotional intonations in both speech conditions it becomes apparent that prosodic speech indeed engages the left temporal (mSTR, pSTR) regions. Under the assumption that a determining parameter of prosody, namely pitch, is lateralized in the right hemisphere (e.g., Meyer et al., 2002; Tzourio et al., 1997; Zatorre & Belin, 2001; Zatorre & Samson, 1991), the reversed lateralization of the STR activation in the current experiment might be a reflection of the task at

hand which forces the listener to search for meaning in degraded speech in order to make a judgment. Most importantly, this effect occurs independent of the valence of the emotional contour.

Two issues remain. First, the activation pattern correlated with the perception of emotional prosody was bilateral, but left accentuated in particular in the prosodic speech condition. Second, we did not find any clear lateralization of valence (i.e., negative vs. positive emotional intonations).

Given the few fMRI (Buchanan et al., 2000; George et al., 1996) studies and a vast majority of lesion studies (e.g., Blonder et al., 1991; Starkstein et al., 1994) that report right hemisphere lateralization of emotional prosody, our results deserve a solid discussion. Regarding the first issue there is lesion evidence that strongly supports the notion of a bilateral distribution of emotional prosody (Cancelliere & Kertesz, 1990; Van Lancker & Sidtis, 1992) and this is in line with the current data. However, two brain potential studies (Kotz et al., 2000; Pihan et al., 2000) suggest that the lateralization of emotional prosody can vary as a function of task demands, a conclusion which is also supported by a patient study. Tompkins and Flowers (1985) reported that with increasing cognitive complexity of a task, patients with left hemisphere lesions showed a similar breakdown in the evaluation of semantically neutral, but emotionally intoned sentences, as patients with right hemisphere lesions. Thus, activation in the left hemisphere (frontal and temporal) might correlate with increasing cognitive demands when tasks involve complex processes. In line with this argument, two conclusions can be drawn from the current data. The perception of emotional prosody correlates with bilateral activation, thus calling into question a pure right hemisphere lateralization hypothesis of emotional prosody. A shift to a left-accentuated but bilateral pattern of activation in prosodic speech indicates that task-specific processes (i.e., type and demand) influence the perception of emotional prosody. In addition, to our knowledge this is the first imaging study on the lateralization of emotional prosody that used a single-trial pseudorandomized presentation of both the normal and prosodic speech conditions. So, while presenting many advantages, switching between the two speech conditions may have created a possible influence of normal speech on prosodic speech.

Regarding the issue of emotional valence, previous imaging studies by George et al. (1996) and Buchanan et al. (2000) do not report lateralization of emotional prosody as a function of valence. This null result is further supported by some recent lesion data that correlated aprosodic syndromes with bilateral lesion sites (Cancelliere & Kertesz, 1990; Ross et al., 1997). Thus, the question remains to explain the discrepancies presented in the experimental literature on emotional prosody.

5. Conclusion

In summary, the current data support the notion that an exclusive right hemispheric lateralization of emotional prosody is not justified. Functionally, the present data together with those in the literature suggest that both hemispheres engage in the perception of emotional prosody, but that the left hemisphere can play a stronger role with increasing task demands. The data, moreover indicate that lateralization does not necessarily vary as a function of emotional valence. Methodologically, the use of prosodic speech clearly helps to specify which areas of the brain recruit bilateral activation during the perception of emotional prosody. The combined findings suggest that both frontal and temporal regions, but most importantly the basal ganglia play a functional role in emotional speech processing.

Acknowledgments

The authors would like to thank Kerrie Elston-Guettler for helpful comments on an earlier version of the paper and Sonja Latner and Andrea Gast-Sandmann for graphics support.

References

- Baum, S. R., & Pell, M. D. (1999). The neural basis of prosody: Insight from lesion studies and neuroimaging. *Aphasiology*, *13*, 581–608.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Springer, J. A., Kaufmann, J. N., & Possing, E. T. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex*, *10*, 512–528.
- Blonder, L. X., Bowers, D., & Heilman, K. M. (1991). The role of the right hemisphere in emotional communication. *Brain*, *114*, 1115–1127.
- Blonder, L. X., Gur, R. E., & Gur, R. C. (1989). The effects of right and left hemiparkinsonism on prosody. *Brain and Language*, *36*, 193–207.
- Borod, J. C. (1993). Cerebral mechanisms underlying facial, prosodic, and lexical emotional expression: A review of neuropsychological studies and methodological issues. *Neuropsychology*, *7*, 445–463.
- Bosch, V. (2000). Statistical analysis of multi-subject fMRI data: The assessment of focal activations. *Journal of Magnetic Resonance Imaging*, *11*, 61–64.
- Brådvik, B., Dravins, C., Holtås, Rosen, I., Ryding, E., & Ingvar, D. (1991). Disturbances of speech prosody following right hemisphere infarcts. *Acta Neurologica Scandinavica*, *84*, 114–126.
- Bryan, K. (1989). Language prosody in the right hemisphere. *Aphasiology*, *3*, 285–299.
- Buchanan, T. W., Lutz, K., Mirsazade, S., Specht, K., Shah, N. J., Zilles, K., & Jäncke, L. (2000). Recognition of emotional speech prosody and verbal components of spoken language: An fMRI study. *Cognitive Brain Research*, *9*, 227–238.
- Burton, M. W., Small, S. L., & Blumstein, S. E. (2000). The role of segmentation in phonological processing: An fMRI investigation. *Journal of Cognitive Neuroscience*, *12*(4), 679–690.
- Cancelliere, A., & Hausdorf, P. (1987). *Emotional expression and comprehension in Huntington's disease*. Poster presented at the

- International Neurophysiological Society Annual Meeting, New Orleans, LA.
- Cancelliere, A., & Kertesz, A. (1990). Lesion localization in acquired deficits of emotional expression and comprehension. *Brain and Cognition*, 13, 133–147.
- Caplan, D., Alpert, N., & Waters, G. (1999). PET studies of syntactic processing with auditory sentence presentation. *NeuroImage*, 9, 343–351.
- Carpentier, A., Pugh, K. R., Westerveld, M., Studholme, C., Skinjar, O., Thompson, J. L., Spencer, D. D., & Constable, R. T. (2001). Functional MRI of language processing: Dependence on input modality and temporal lobe epilepsy. *Epilepsia*, 42(10), 1241–1254.
- Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, 40(2), 141–201.
- Dapretto, M., & Bookheimer, S. Y. (1999). Form and content: Dissociating syntax and semantics in sentence comprehension. *Neuron*, 24, 427–432.
- Davidson, R. J., Abercrombie, H., Nitschke, J. B., & Putnam, K. (1999). Regional brain function emotion and disorders of emotion. *Current Opinion in Neurobiology*, 9, 228–234.
- Dehaene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, E., Perani, D., van de Moortele, P., Lehericy, S., & Le Bihan, D. (1997). Anatomical variability in the cortical representation of first and second language. *NeuroReport*, 8, 3809–3815.
- Demonet, J. F., Price, C., Wise, R., & Frackowiak, R. S. (1994). A PET study of cognitive strategies in normal subjects during language tasks. Influence of phonetic ambiguity and sequence processing on phoneme monitoring. *Brain*, 117(4), 671–682.
- Dykstra, K., Gandour, J., & Stark, R. E. (1995). Disruption of prosody after frontal lobe seizures in the non-dominant hemisphere. *Aphasiology*, 9(5), 453–476.
- Emmorey, K. (1987). The neurological substrates for prosodic aspects of speech. *Brain and Language*, 30, 305–320.
- Friederici, A. D., Meyer, M., & von Cramon, D. Y. (2000). Auditory language comprehension: An event-related fMRI study on the processing of syntactic and lexical information. *Brain and Language*, 74, 289–300.
- Friston, K. J. (1994a). Statistical parametric mapping. In R. W. Thatcher, M. Hallet, T. Zeffiro, E. R. John, & M. Huerta (Eds.), *Functional neuroimaging* (pp. 79–93). San Diego: Academic Press.
- Friston, K. J. (1994b). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2, 189–210.
- Friston, K. J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M. D., & Turner, R. (1998). Event-related fMRI: Characterizing differential responses. *NeuroImage*, 7, 30–40.
- Friston, K. J., Holmes, A. P., Poline, J.-B., Grasby, B. J., Williams, C. R., Frackowiak, R. S. J., & Turner, R. (1995a). Analysis of fMRI time-series revisited. *NeuroImage*, 2, 45–53.
- Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., & Frackowiak, R. S. J. (1995b). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2, 189–210.
- George, M. S., Parekh, P. I., Rosinsky, N., Ketter, T. A., Kimbell, T. A., Heilmann, K. M., Herscovitch, M. D., & Post, R. M. (1996). Understanding emotional prosody activates right hemisphere regions. *Archives of Neurology*, 53, 665–670.
- Giraud, A. L., Kell, C., Klinke, R., Russ, M. O., Sterzer, P., Thierfelder, C., Preibisch, C., & Kleinschmidt, A. (2001). Dissociating effort and success in speech comprehension: An fMRI study. *NeuroImage*, 13, S534.
- Hall, D. A., Haggard, M. P., Akeroyd, M. A., Summerfield, A. Q., Palmer, A. R., Elliot, M. R., & Bowtell, R. W. (2000). Modulation and task effects in auditory processing measured using fMRI. *Human Brain Mapping*, 10, 107–119.
- Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Science*, 4, 131–138.
- Indefrey, P., Haggart, P., Herzog, H., Seitz, R. J., & Brown, C. (2001). Syntactic processing in left prefrontal cortex is independent of lexical meaning. *NeuroImage*, 14, 546–555.
- Josephs, O., Turner, R., & Friston, K. (1997). Event-related fMRI. *Human Brain Mapping*, 5, 243–248.
- Kotz, S. A., Alter, K., Besson, M., Schirmer, A., & Friederici, A. D. (2000). The interface between prosodic and semantic processes: An ERP study. *Supplement of the Journal of Cognitive Neuroscience*, 121.
- Lee, J. H., Garwood, M., Menon, P., Adriany, G., Andersen, P., Truwit, C. L., & Ugurbil, K. (1995). High contrast and fast three-dimensional magnetic resonance imaging at high fields. *Magnetic Resonance in Medicine*, 34, 308.
- Lohmann, G., Mueller, K., Bosch, V., Mentzel, H., Hessler, S., Chen, L., & von Cramon, D. Y. (2001). Lipsia—A new software system for the evaluation of functional magnetic resonance images of the human brain. *Computerized Medical Imaging and Graphics*, 25, 449–457.
- Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Music syntax is processed in the area of Broca: An MEG study. *Nature Neuroscience*, 4, 540–545.
- Meyer, M., Alter, K., Friederici, A. D., Lohmann, G., & von Cramon, D. Y. (2002). Functional MRI evidence reveals brain regions mediating slow prosodic modulations in spoken sentences. *Human Brain Mapping*, 17, 73–88.
- Meyer, M., Alter, K., Steinhauer, K., & Friederici, A. D. (2001). Cerebral substrates of pitch modulations in sentences: Evidence from 3T fMRI. *Journal of Cognitive Neuroscience* (Suppl.), 158.
- Meyer, M., Friederici, A. D., & von Cramon, D. Y. (2000). Neurocognition of auditory sentence comprehension: Event-related fMRI reveals sensitivity to syntactic violations and task demands. *Cognitive Brain Research*, 9, 19–33.
- Nauta, H. J. W. (1986). The relationship of the basal ganglia to the limbic system. In P. J. Vinken, G. W. Bruyn, & H. L. Klawans (Eds.), *Handbook of clinical neurology* (Vol. 49). New York: F.A. Davis.
- Ouellette, G., & Baum, S. R. (1993). Acoustic analysis of prosodic cues in left and right-hemisphere-damaged patients. *Aphasiology*, 8(3), 257–283.
- Pell, M. D. (1998). Recognition of prosody following unilateral brain lesion: influence of functional and structural attributes of prosodic intonations. *Neuropsychologia*, 36, 701–715.
- Pell, M. D., & Baum, S. R. (1997). The ability to perceive and comprehend intonation in linguistic and affective contexts by brain-damaged adults. *Brain and Language*, 57, 80–89.
- Pihan, H., Ackermann, H., & Altenmüller, E. (1997). The cortical processing of perceived emotion: A DC potential study on affective prosody. *Neuroreport*, 8, 623–627.
- Pihan, H., Altenmüller, E., Hertrich, I., & Ackermann, H. (2000). Cortical activation patterns of affective speech processing depend on concurrent demands on the subvocal rehearsal system: A dc-potential study. *Brain*, 123, 2338–2349.
- Poldrack, R. A., Temple, E., Protopapas, A., Nagarajan, S., Tallal, P., Merzenich, M., & Gabrieli, J. D. E. (2001). Relations between neural bases of dynamic auditory processing and phonological processing: Evidence from fMRI. *Journal of Cognitive Neuroscience*, 13, 687–697.
- Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *NeuroImage*, 10, 15–35.
- Riecker, A., Ackermann, H., Wildgruber, D., Dogil, G., & Grodd, W. (2000). Opposite hemispheric lateralization during speaking and singing at motor cortex, insula and cerebellum. *NeuroReport*, 11, 1997–2000.

- Ross, E. D. (1981). The aprosodias: Functional-anatomic organization of the affective components of language in the right hemisphere. *Archives of Neurology*, *38*, 561–569.
- Ross, E., D., Edmondson, J. A., Seibert, G. B., & Homan, R. W. (1988). Acoustic analysis of affective prosody during right-sided Wada test: A within-subjects verification of the right hemisphere's role in language. *Brain and Language*, *33*, 128–145.
- Ross, E. D., Thompson, R. D., & Yenkosky, J. (1997). Lateralization of affective prosody in the brain and the callosal integration of hemispheric language functions. *Brain and Language*, *56*, 27–54.
- Scott, S., Caird, F., & Williams, B. (1984). Evidence for an apparent sensory speech disorder in Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, *47*, 840–843.
- Sonntag, G. P., & Portele, T. (1998). PURR—a method for prosody evaluation and investigation. *Journal of Computer Speech and Language*, *12*, 437–451.
- Starkstein, S. E., Federoff, J. P., Price, R. C., Leiguarda, R. C., & Robinson, R. G. (1994). Neuropsychological and neuroradiological correlates of emotional prosody comprehension. *Neurology*, *44*, 515–522.
- Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience*, *2*(2), 191–196.
- Talairach, P., & Tournoux, J. (1988). *A stereotactic coplanar atlas of the human brain*. Stuttgart: Thieme.
- Tompkins, C. A., & Flowers, C. R. (1985). Perception of emotional intonation by brain-damaged adults: The influence of task processing levels. *Journal of Speech and Hearing Research*, *28*, 527–538.
- Tzourio, N., Massiou, F. E., 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, S. G., Menon, R., Merkle, H., & Ogawa, S. R. S. (1993). Magnetic fields: Initial experiences at 4T. *Magnetic Resonance Quarterly*, *9*, 259.
- Van Lancker, D. (1980). Cerebral lateralization of pitch cues in the linguistic signal. *International Journal of Human Communication*, *13*(2), 101–109.
- Van Lancker, D., & Sidtis, J. J. (1992). The identification of affective-prosodic stimuli by left- and right-hemisphere-damaged subjects: All errors are not equal. *Journal of Speech and Hearing*, *35*, 963–970.
- Warren, P., Grabe, E., & Nolan, F. (1995). Prosody, phonology, and parsing in closure ambiguities. *Language and Cognitive Processes*, *10*, 457–486.
- Weintraub, S., Mesulam, M.-M., & Kramer, L. (1981). Disturbances in prosody: A right-hemisphere contribution to language. *Archives of Neurology*, *38*, 742–744.
- Worsley, K. J., & Friston, K. J. (1995). Analysis of fMRI time-series revisited—again. *NeuroImage*, *2*, 359–365.
- Zarahn, E., Aguirre, G. K., & D'Esposito, M. (1997). Empirical analyses of BOLD fMRI statistics. *NeuroImage*, *5*, 179–197.
- Zatorre, R. J. (1988). Pitch perception of complex tones and human temporal-lobe function. *Journal of the Acoustical Society of America*, *84*, 566–572.
- Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, *11*, 946–953.
- Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Science*, *1*(6), 37–46.
- Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science*, *256*, 846–849.
- Zatorre, R. J., & Samson, S. (1991). Role of the right temporal neocortex in retention of pitch in auditory short-term memory. *Brain*, *114*, 2403–2417.