

## Identification of emotional intonation evaluated by fMRI

D. Wildgruber,<sup>a,b,\*</sup> A. Riecker,<sup>a,b</sup> I. Hertrich,<sup>a</sup> M. Erb,<sup>b</sup> W. Grodd,<sup>b</sup>  
T. Ethofer,<sup>a</sup> and H. Ackermann<sup>a</sup>

<sup>a</sup>Department of General Neurology, Hertie Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany

<sup>b</sup>Section MR of CNS, Department of Neuroradiology, University of Tübingen, Germany

Received 19 July 2004; revised 15 October 2004; accepted 28 October 2004  
Available online 15 December 2004

During acoustic communication among human beings, emotional information can be expressed both by the propositional content of verbal utterances and by the modulation of speech melody (affective prosody). It is well established that linguistic processing is bound predominantly to the left hemisphere of the brain. By contrast, the encoding of emotional intonation has been assumed to depend specifically upon right-sided cerebral structures. However, prior clinical and functional imaging studies yielded discrepant data with respect to interhemispheric lateralization and intrahemispheric localization of brain regions contributing to processing of affective prosody. In order to delineate the cerebral network engaged in the perception of emotional tone, functional magnetic resonance imaging (fMRI) was performed during recognition of prosodic expressions of five different basic emotions (happy, sad, angry, fearful, and disgusted) and during phonetic monitoring of the same stimuli. As compared to baseline at rest, both tasks yielded widespread bilateral hemodynamic responses within frontal, temporal, and parietal areas, the thalamus, and the cerebellum. A comparison of the respective activation maps, however, revealed comprehension of affective prosody to be bound to a distinct right-hemisphere pattern of activation, encompassing posterior superior temporal sulcus (Brodmann Area [BA] 22), dorsolateral (BA 44/45), and orbitobasal (BA 47) frontal areas. Activation within left-sided speech areas, in contrast, was observed during the phonetic task. These findings indicate that partially distinct cerebral networks subservise processing of phonetic and intonational information during speech perception.

© 2004 Elsevier Inc. All rights reserved.

**Keywords:** Affect; Communication; Emotion; fMRI; Language; Lateralization; Prosody; Phonetics; Valence

### Introduction

During speech production, information about a speaker's emotional state is predominantly conveyed by the modulation of intonation (affective prosody). At the perceptual level, emotional tone is characterized by variations of pitch, syllable durations, loudness, and voice quality across utterances (suprasegmental features) imposed upon segmental verbal information encoded in phonetic/phonological units (Ackermann et al., 1993; Bachorowski and Owren, 2003; Banse and Scherer, 1996; Cutler et al., 1997; Sidtis and Van-Lancker-Sidtis, 2003). As concerning cerebral topography of prosody processing, observations in patients suffering from focal brain lesions indicate that the well-established left-sided dominance for language comprehension does not extend to perception of emotional intonation (Adolphs, 2002; Baum and Pell, 1999; Borod et al., 2001, 2002; Charbonneau et al., 2003; Pell and Baum, 1997). According to the neuroanatomical model proposed by Ross (1981), prosodic information is processed within distinct right-sided perisylvian regions that are organized in complete analogy to left-sided language areas. Expression of affective prosody, thus, is believed to rely on the Broca-homotop within the right inferior frontal cortex, whereas comprehension of intonational information is presumed to be bound to the Wernicke-homotop within the right superior temporal region. However, empirical evidence for this model as provided by Ross (1981) was based on a few case reports only, and more systematic investigations yielded divergent results. Nevertheless, as concerns comprehension of speech melodies, the findings of the majority of lesion studies seem to be compatible with the assumption that perceptual prosodic functions are predominantly bound to the right posterior perisylvian cortex (Borod et al., 2002; Darby, 1993; Heilman et al., 1984; Starkstein et al., 1994). In addition, various clinical examinations indicate a widespread network of—partially bilateral—cerebral regions including the frontoparietal cortex (Adolphs et al., 2002; Breitenstein et al., 1998) and the basal ganglia (Breitenstein et al., 1998, 2001; Cancelliere and Kertesz, 1990; Pell and Leonard, 2003) to contribute to comprehension of emotional intonation. In line with these findings, neuroimaging studies as a rule yielded rightward lateralization of hemodynamic

---

\* Corresponding author. Department of General Neurology, Hertie Institute for Clinical Brain Research, University of Tübingen, Hoppe-Seyler-Str. 3, 72076 Tübingen, Germany. Fax: +49 7071 294371.

E-mail address: dirk.wildgruber@med.uni-tuebingen.de (D. Wildgruber).

Available online on ScienceDirect (www.sciencedirect.com).

activation within posterior temporal regions (Buchanan et al., 2000; Kotz et al., 2003; Mitchell et al., 2003; Wildgruber et al., 2002) and revealed additional—partially bilateral—responses within the frontal cortex (Buchanan et al., 2000; George et al., 1996; Imaizumi et al., 1997; Kotz et al., 2003; Wildgruber et al., 2002, 2004), the anterior insula (Imaizumi et al., 1997; Wildgruber et al., 2002, 2004), and the basal ganglia (Kotz et al., 2003) during recognition of affective intonation. Considerable differences in lateralization and exact localization of cerebral responses, however, did not allow for an indisputable identification of the neural substrate of prosody processing. Presumably, the observed discrepancies are due to differences in methodology including stimulus selection, task conditions, control conditions, and imaging modalities. It has been proposed, for example, that extraction of specific acoustic properties underlying emotional prosody relies on differentially lateralized cerebral regions (Ackermann et al., 2001; Sidtis and Van-Lancker-Sidtis, 2003; Van Lancker and Sidtis, 1992; Wildgruber et al., 2002; Zatorre, 2001). Moreover, the selection of emotional categories included in the stimulus material might matter, since the valence of emotional expression has been reported to influence lateralization of cerebral responses (Canli et al., 1998; Davidson and Tomarken, 1989; Davidson et al., 1999; Murphy et al., 2003). As concerns speech intonation, several clinical examinations failed to show any interactions between hemispheric lateralization and emotional valence (Baum and Pell, 1999; Borod et al., 2002; Kucharska-Pietura et al., 2003; Pell, 1998). Considering functional imaging data, however, distinct cerebral activation patterns bound to specific emotional categories such as disgust, anger, fear, or sadness have been observed during perception of facial emotional expressions (Kesler-West et al., 2001; Murphy et al., 2003; Phan et al., 2002; Sprengelmeyer et al., 1998). Several studies have corroborated the notion that responses of the amygdala are specifically related to facial expressions of fear (Adolphs, 2002; Phan et al., 2002), whereas facial expressions of disgust seem to be linked to activation of the anterior insula (Calder et al., 2000; Phan et al., 2002; Phillips et al., 1998; Sprengelmeyer et al., 1998; Wicker et al., 2003). As concerns vocal emotional expressions, fear-specific responses within the amygdalae have also been reported (Morris et al., 1999; Phillips et al., 1998), whereas the predicted disgust-related activation of the anterior insula has not been observed in a prior PET experiment (Phillips et al., 1998). It is unsettled, thus, to which extent lateralization and exact localization of cerebral activation during comprehension of emotional prosody is linked to specific emotional categories. Based on the aforementioned clinical and neuroimaging studies, presumably there are cerebral regions—including the right posterior temporal cortex—that contribute to comprehension of emotional prosody independent of any specific emotional content, whereas other regions—including the amygdala and anterior insula—are selectively linked to comprehension of specific emotional categories.

To allow for a separation of these components, the stimulus material used in the present study comprised sentences spoken in emotional intonation of five different basic emotions (happiness, anger, fear, sadness, and disgust). Verbal utterances were presented to healthy subjects under two different task conditions during functional magnetic resonance imaging (fMRI). On one hand, subjects were asked to identify the emotion expressed by the tone of voice, on the other, they performed a phonetic identification task. Since both tasks require evaluation of completely identical acoustic stimuli and involve very similar response mechanisms, comparison

of the respective hemodynamic activation patterns allows for the separation of task-specific cerebral responses independently of stimulus characteristics and unspecific task components. In order to delineate cerebral structures contributing to identification of affective prosody independent of specific emotional categories, responses during identification of affective prosody across all emotional categories were compared to the phonetic control condition. To disentangle patterns of cerebral activation related to comprehension of specific emotional categories, each emotional category was compared against the others. The main goal of the study, thus, was to test predictions for the following two questions:

- (a) Which areas of the human brain contribute to identification of affective intonation independent of specific emotional information conveyed? Hypothesis: a network of right-sided regions including the right posterior temporal cortex.
- (b) Does the localization of responses vary for specific emotions? Hypothesis: perception of different emotional categories is associated to specific brain regions, such as fear-specific responses within the amygdalae and disgust-specific responses within the anterior insula.

## Materials and methods

In order to generate appropriate verbal stimuli, 100 short German declarative sentences with emotionally neutral content were selected, such as “Der Gast hat sich für Donnerstag ein Zimmer reserviert [The visitor reserved a room for Thursday]”, “Die Anrufe werden automatisch beantwortet [Phone calls are answered automatically]”. Sentences were randomly ascribed to one of five different target emotions (happiness, anger, fear, sadness, or disgust) and spoken by a professional actress and an actor, expressing the respective emotion by modulation of affective intonation. Duration of verbal utterances ranged from 2.2 to 3.3 s. In order to select the stimuli that most clearly represent the intended affective prosody, tape recordings of all 200 utterances (20 sentences  $\times$  5 emotions  $\times$  2 actors) were presented to 10 healthy subjects (five males, five females, age 22–28 years). During this pretest, the acoustic stimuli were presented together with a tape recording of scanner noise to simulate the acoustic environment of an fMRI experiment. Under these conditions, identification of emotional intonation across all 200 sentences was performed at an average accuracy of 90.3%. There were, however, some differences between the emotional categories. Happy, fearful, angry, and sad intonations were identified with accuracy rates ranging from 90% to 94%, whereas only 77% of expressions of disgust could be identified correctly. Five sentences per emotional category that had been identified with high accuracy were selected for the fMRI experiment. Post hoc analysis of this set of 50 stimuli (5 sentences  $\times$  5 emotions  $\times$  2 speakers) yielded comparable accuracy rates ranging from 91% to 94% across the different emotional categories (mean:  $92.3 \pm 5.0\%$ ).

Ten different right-handed subjects (five males, five females, age 21–33 years) without history of neurological or psychiatric diseases participated in the fMRI experiment. Informed consent was obtained according to the Declaration of Helsinki. The Ethical Committee of the University of Tuebingen had approved the investigation. Handedness was determined using the Edinburgh Inventory (Oldfield, 1971). All 10 participants showed a lateralization quotient of 100%.

Two different tasks were carried out in randomized order during six measurement sessions. Each session of the experiment comprised acquisition of 110 functional images during a temporal interval of 5 min 30 s. The set of 50 stimuli was presented twice, once under each of the task conditions. During the emotion identification task, subjects were asked to name the emotion expressed by the tone of voice, whereas the phonetic task required verbalization of the vowel following the first /a/ in each sentence. In analogy to the emotion identification task, vowel identification also represented a forced choice selection from five distinct vowels (/a/, /e/, /i/, /o/, /u/). In both tasks, subjects were asked to give a verbal response as quickly as possible and they were given a list of possible response alternatives prior to testing. All sentences comprised 14 syllables and the target vowel appeared at different positions within the sentence (on average after 3.9 syllables with a standard deviation of 1.9 syllables). Stimuli were presented asynchronously in pseudorandomized order. Interstimulus intervals ranged from 12.6 to 25.2 s. Presentation of acoustic stimuli was carried out from a PC via a dynamic fMRI-adapted headphone system (Baumgart et al., 1998). Subjects' verbal responses to each stimulus were digitally recorded on a PC using an fMRI-adapted microphone (Price et al., 2001) positioned close to the mouth inside the head coil to allow for the evaluation of behavioral data during the fMRI experiment (Fig. 1).

Participants lay supine in a 1.5-T whole body scanner (Siemens Vision), their eyes closed and their heads supported by a foam rubber within the head coil. fMRI data were acquired continuously using a multislice EPI sequence (Klose et al., 1999) covering 28 parallel axial slices (4-mm thickness, 1-mm gap, TR 3 s, TE 39 ms,  $\alpha$  90°, FOV 192 mm, 64<sup>2</sup> matrix). High-resolution images obtained with a T1-weighted 3D-Turbo-Flash-Sequence served as an anatomical reference. The first five fMRI images acquired during each measurement session were discarded from further analysis in order to exclude measurements preceding T1 equilibrium. Post-processing of functional images including 3D motion correction, slice time correction, normalization into MNI space (Montreal Neurological Institute, Collins et al., 1994), and spatial smoothing

(10 mm FWHM) relied on SPM99 (Wellcome Department of Cognitive Neurology, London, UK) (<http://www.fil.ion.ucl.ac.uk/spm>). The evoked hemodynamic responses were modeled as delta functions convolved with a synthetic hemodynamic response function and its temporal and spatial derivative within the context of the general linear model. All events were time locked to the onset of the acoustic stimuli. Statistical evaluation of group data was based on second-level random-effect analysis (height threshold:  $T > 4.30$ , corrected at cluster level,  $P < 0.05$ , extent threshold:  $k > 38$  voxels). Three different contrasts were calculated:

- Main effects (activation vs. rest): to determine the whole network of brain region participating in the processing of acoustic stimuli under the two different task conditions, the responses obtained during the emotional and phonological identification task were each compared to the respective baseline at rest.
- Task-specific responses (emotion identification vs. vowel identification): to identify brain areas exhibiting hemodynamic responses specifically related either to comprehension of emotional intonation or to phonetic processing, the respective activation maps were compared.
- Influence of specific emotions (ANOVA across emotional categories): data obtained during the emotion identification condition were further analyzed for significant category-specific activation patterns. To this end, perception of stimuli expressing each emotional category (happy, fearful, sad, angry, disgusted) was contrasted against the four other categories (ANOVA design and multiple  $t$  tests).

## Results

During the fMRI experiment, subjects correctly identified sentences with respect to emotional tone at a slightly lower rate (mean:  $75.2 \pm 7.9\%$ ) as compared to vowel identification (mean:  $83.4 \pm 7.0\%$ ,  $P < 0.05$ ). The accuracy scores for happy (90%),

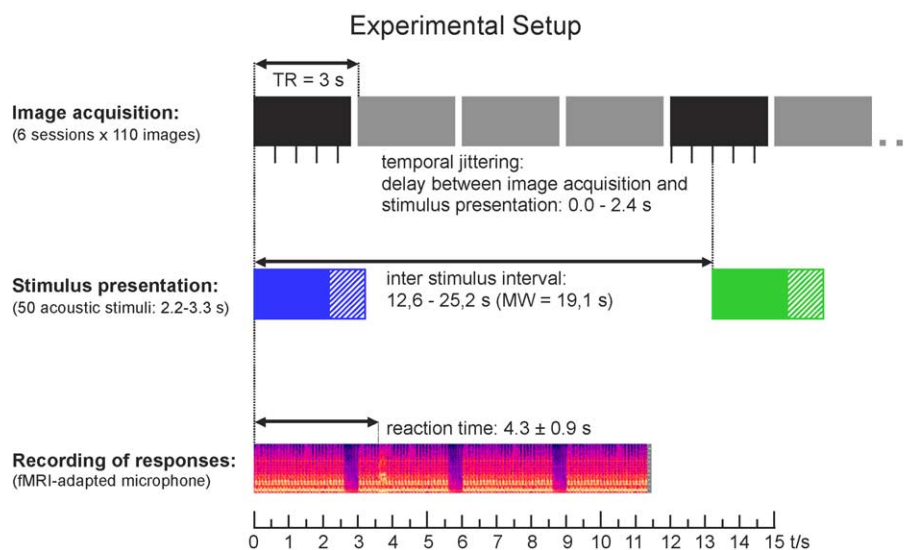


Fig. 1. Experimental setup. The emotional and the phonological identification tasks were carried out in randomized order across six imaging sessions comprising acquisition of 110 EPI images each (upper row). Acoustic stimuli were presented in intervals ranging from 12.6 to 25.2 s via MR-adapted headphones (middle row). Verbal responses to the identification tasks were recorded with a MR-adapted microphone to analyze reaction times and accuracy rates during the fMRI experiment (lower row).

angry (82%), and sad (84%) expressions did not differ significantly from pretest ratings, whereas fearful (51%) and disgusted (57%) expressions were identified at significantly lower rates ( $P < 0.05$ ) (Fig. 2a). Response times for the emotional task (mean:  $4.3 \pm 0.9$  s) showed no significant differences as compared to the phonetic task (mean:  $4.1 \pm 1.0$  s). Mean values of response times ranged from 3.8 s (happy) to 5.2 s (disgusted) across the different emotional categories (Fig. 2b).

Cerebral responses obtained during both tasks, as compared to the rest condition, yielded a bilateral network of hemodynamic activation at the level of cortical and subcortical regions including frontal, temporal and parietal cortex, thalamus, and cerebellum (Fig. 3, Table 1). To identify brain regions specifically contributing to the processing of emotional intonation, the respective activation patterns were directly compared to the responses obtained during phonetic processing of the identical acoustic stimuli. Using this subtraction approach, responses within two activation clusters, localized within the right superior temporal sulcus (BA 22/42) and the right inferior frontal cortex (BA 45/47), could be ascribed to emotional identification (Fig. 4, Table 1). Individual signal time curves demonstrated hemodynamic activation within the right superior temporal sulcus (BA 22/42) under both task conditions in all subjects with significantly higher amplitudes during the emotional task. In contrast, stimulus-related activation within the right orbitobasal frontal cortex (BA 47) was limited to the emotional task (Fig. 5). The phonetic task yielded activation within the left dorsolateral frontal region (BA 9/44) and bilateral responses within the inferior parietal lobule (BA 7/40). Evaluation of responses bound to distinct emotional categories did not show significant activation differences for any specific emotional intonation as compared to the others.

## Discussion

Evaluation of the behavioral data revealed considerable differences in accuracy rates across emotional categories. Lower identification rates for prosodic expression of disgust and fear as compared to happy, sad, and angry vocalizations are in good accordance with prior observations and might be related to

differential degrees of acoustic recognizability for specific emotions (Banse and Scherer, 1996). Significantly less accurate performance on these emotional expressions during the fMRI experiment as compared to the pretest, however, was unexpected and might indicate a ceiling effect of emotional identification for all emotional categories during the pretest. Most presumably, an unspecific increase of task difficulty caused by distraction of attention within the scanner environment might have unmasked different degrees of emotional expressiveness or intelligibility.

Longer reaction times during evaluation of disgusted expressions as compared to the other emotional categories, moreover, support the assumption of higher demands on cerebral processing capacities during evaluation of these stimuli. Similar reaction times across the emotional and the phonetic identification task, however, seem to indicate comparable levels of overall task difficulty.

### Network of brain regions contributing to emotional and phonetic identification

During perception of each sentence, continuous extraction and encoding of the acoustic features of the speech signal as well as transient storage and processing of the respective data are required for evaluation of emotional intonation as well as vowel identification. Since verbal utterances were presented during both tasks, phonetic, syntactic, and semantic encoding must be expected, even though higher level linguistic processing was not explicitly required. Moreover, emotional intonation of verbal utterances must be expected to induce an involuntary analysis of speech prosody to some degree. Finally, verbal responses—requiring speech motor control mechanisms—were used to indicate a subject's decision on the respective identification task. Both tasks, thus, required several common cognitive and motor operations. Therefore, the observed high degree of similarity in hemodynamic response patterns is not surprising.

Activation within the superior temporal cortex and the thalamus indicates engagement of the primary and secondary auditory cortices and the respective relay station of the acoustic pathway. Responses within the motor system—including bilateral supplementary motor cortex, rolandic region, and cerebellum—

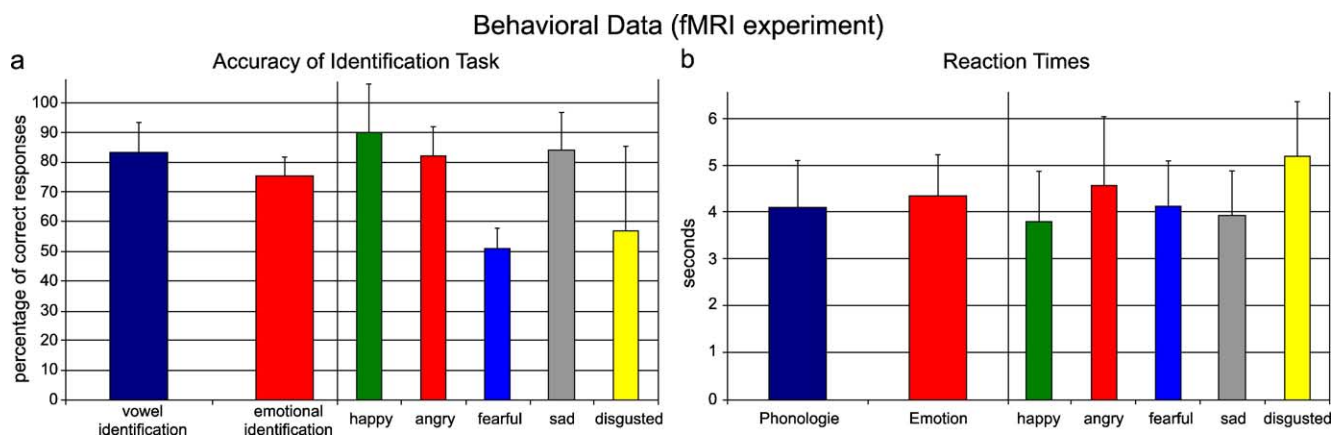


Fig. 2. Behavioral data obtained during the fMRI experiment. The percentage of correct answers (a) was significantly lower during identification of emotional intonation (mean:  $75.2 \pm 7.9\%$ ) as compared to vowel identifications (mean:  $83.4 \pm 7.0\%$ ,  $P < 0.05$ ). Recognition rates for specific emotions ranged between 51% (fear) and 92% (happiness). Response times, taken from the onset of acoustic stimuli to the onset of verbal responses (b), did not show significant differences between the emotional and the phonological task.

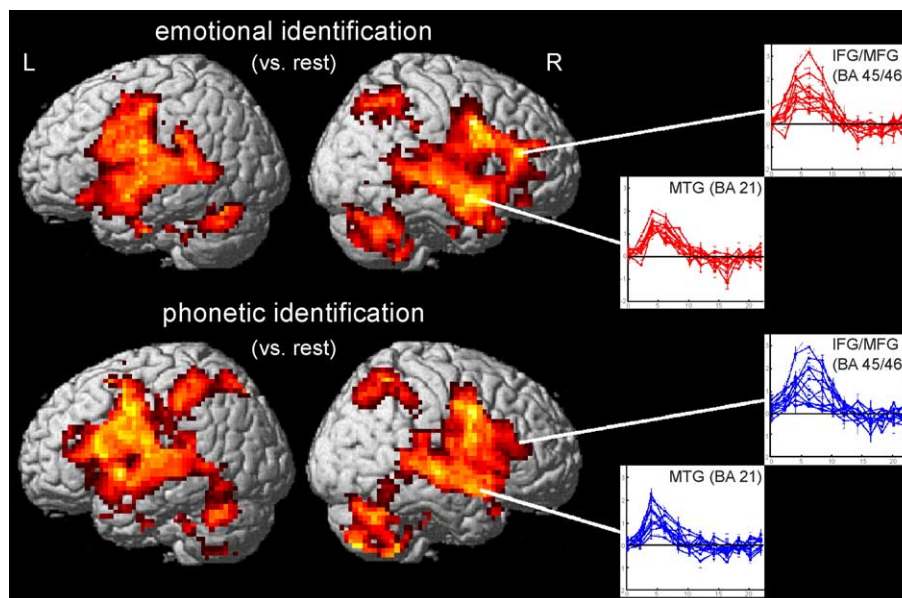


Fig. 3. Cerebral activation during the emotional and phonological identification task compared to rest. Significantly activated areas are projected upon the cortical surface of a template brain with color-coded  $T$  values from red (low) to yellow (high). Signal time courses (right) at the two highest activated voxels during judgement of emotional intonation are shown for all subjects under both task conditions (red = emotional identification; blue = phonological identification). IFG = inferior frontal gyrus; MFG = middle frontal gyrus; MTG = middle temporal gyrus. SPM99, random-effect analysis,  $T > 4.30$ ,  $k > 38$ ,  $P < 0.05$ , corrected.

presumably reflect activation of speech motor control mechanisms during verbal responses to the identification tasks after each acoustic stimulus. A partial contribution of these regions to nonmotor aspects common to both tasks, however, cannot be ruled out on the basis of our data. Activation of the inferior parietal lobule (BA 7/40) during both tasks presumably reflects activation of a passive store for phonetic and intonational acoustic information contributing to the working memory system. In line with previous functional imaging studies (Paulesu et al., 1993; Wildgruber et al., 1999; 2002) and clinical observations

(Shallice and Vellar, 1990), the phonological store required during the vowel identification task seems to be lateralized to the left parietal region, whereas the right inferior parietal lobule might contribute to storage of intonational information (Wildgruber et al., 2002; Zatorre, 2001; Zatorre et al., 1992, 1994). Activation of the left dorsolateral inferior frontal region (BA 44) during the vowel identification task might reflect increased demands on the subvocal rehearsal component of the working memory system (Poeppl, 1996) compared to identification of emotional intonation.

Table 1  
Hemodynamic activation during identification of emotions and vowels

		Emotions vs. Baseline	Vowels vs. Baseline	Emotions vs. Vowels	Vowels vs. Emotions
SMA/ACC (BA 6, 24, 32)		9.28 [3,9,57]	11.97 [3,9,51]	–	–
DLFC (BA 9,44–46)	left	9.52 [–48,9,30]	15.14 [–54,6,30]	–	7.18 [–54,6,33]
	right	16.89 [51,33,15]	12.42 [42,0,33]	–	–
OBFC (BA 47)	left	7.67 [–45,30,–12]	–	–	–
	right	7.33 [48,36,–18]	–	8.63 [48, 33, –15]	–
Rolandic area (BA 3/4/6)	left	12.08 [–45,–9,18]	10.11 [–57,0,21]	–	–
	right	13.29 [45,3,24]	10.16 [57,3,18]	–	–
IPL (BA 7,40)	left	–	14.15 [–21,–63,48]	–	10.16 [–48,–39,54]
	right	9.89 [39,–51,48]	15.99 [27,–54,57]	–	10.57 [27,–66,39]
STG (BA 22,41,42)	left	8.96 [–60,–45,9]	8.10 [–60,–30,9]	–	–
	right	12.32 [48,–45,6]	8.18 [51,–42,6]	8.69 [48,–42,3]	–
MTG (BA 21)	left	11.18 [–36,–3,–15]	8.17 [–39,–9,–18]	–	–
	right	16.18 [45,15,–30]	11.68 [54,3,–15]	–	–
Thalamus	left	11.44 [–6,–27,–3]	12.13 [–6,–24,–3]	–	–
	right	15.75 [6,–21,0]	12.89 [9,–21,6]	–	–
Cerebellum	left	13.56 [–30,–60,–27]	7.98 [–33,–54,–33]	–	–
	right	10.86 [42,–63,–33]	9.18 [48,–66,–33]	–	–
	vermis	14.73 [–6,–72,–30]	10.42 [–3,–72,–24]	–	–

$T$  values and MNI coordinates (in square brackets) of highest activated voxels within each region. SMA, supplementary motor cortex; ACC, anterior cingulate cortex; DLFC, dorsolateral frontal cortex; OBFC, orbitobasal frontal cortex; IPL, inferior parietalis lobulus; STG, superior temporal gyrus; MTG, middle temporal gyrus. Respective Brodmann areas (BA) are printed in brackets. SPM99, random-effect analysis,  $n = 10$ ,  $P < 0.05$  corrected,  $T > 4.30$ ,  $k > 38$ .

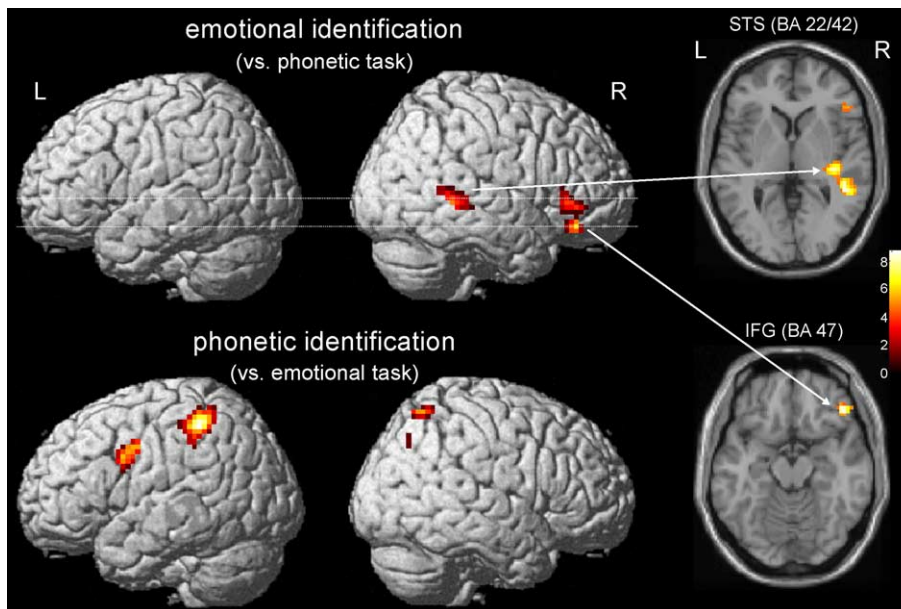


Fig. 4. Emotional identification vs. vowel identification. Significant hemodynamic responses during identification of emotional intonation as compared to vowel identification are superimposed upon the cortical surface of a template brain (upper left row) and upon an axial slice (right) at the level of the highest activated voxels within the activation clusters. The emotional task yielded specific activation within the right superior temporal sulcus (STS, BA 22/42,  $z = 3$  mm) and the right inferior frontal gyrus (IFG, BA 45/47,  $z = -15$  mm). During the phonetic identification task (lower left row), a significant increase of responses was observed within the dorsolateral left inferior frontal cortex (BA 9/44) and bilateral inferior parietal lobules (BA 7/40). SPM99, random-effect analysis,  $T > 4.30$ ,  $k > 38$ ,  $P < 0.05$ , corrected.

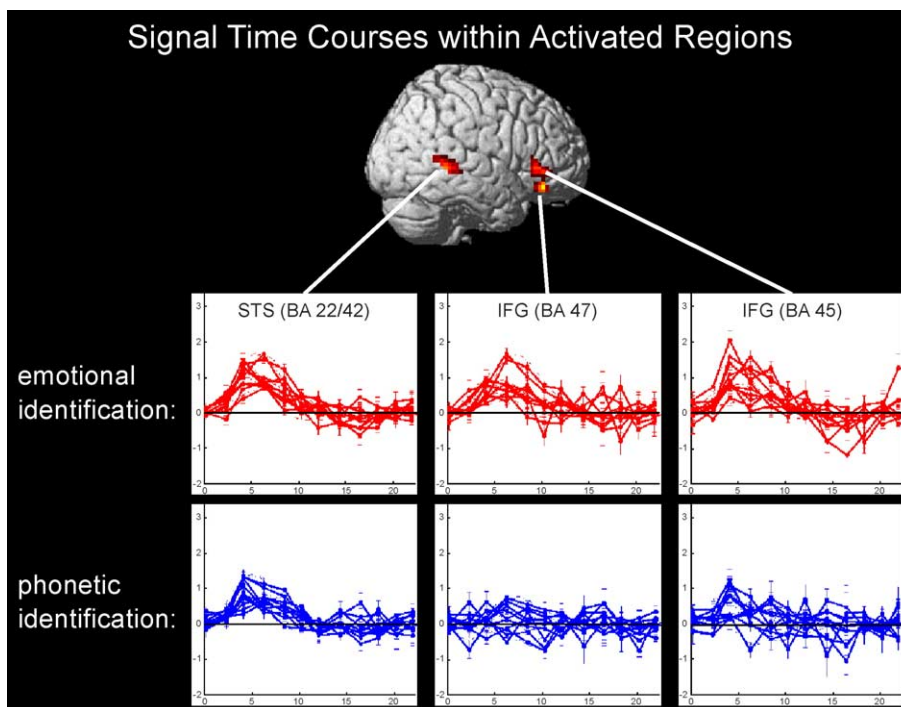


Fig. 5. Stimulus-related signal variation within voxels characterized by local maxima of responses during identification of emotional intonation as compared to the phonological control condition. Signal time curves during the emotional (red) and the phonological (blue) identification task are displayed for all 10 subjects participating to the study. Within the right superior temporal cortex (left panel), a stimulus-related increase of signal intensity was observed under both task conditions in all subjects (with higher amplitudes during the emotional task). Within the right orbitobasal frontal cortex (middle panel), stimulus-related activation is limited to the emotional task, whereas responses within the lateral inferior frontal region (right panel) indicate a slight signal increase during the phonetic task in some subjects. SPM99, random-effect analysis,  $T > 4.30$ ,  $k > 38$ ,  $P < 0.05$ , corrected.

*Hemodynamic responses related to comprehension of emotional intonation independent of emotional categories (emotional identification vs. vowel identification)*

As its main goal, the present study aimed at the identification of brain regions specifically contributing to processing of emotional information conveyed by speech melody. The specificity of hemodynamic responses obtained during experimental conditions depends upon the reference chosen (Newman et al., 2001). Using the resting state of the brain as baseline has been observed to reduce, eliminate, or even reverse the sign of the activity during task conditions (Stark and Squire, 2001). A hierarchical task decomposition attempting to isolate specific processing components of complex tasks, on the other hand, may yield divergent results, depending on slight changes in task-control matching (Newman et al., 2001; Poeppel, 1996). Presentation of completely identical stimuli under two highly specified conditions, therefore, might provide the most appropriate reference condition to evaluate cerebral responses related to specific cognitive operations independent from subsidiary aspects of task performance. In prior imaging studies on emotional prosody, various control conditions have been introduced (i.e., speaker sex judgements, identification of propositional content, rehearsal of the second word). In the present study, the vowel identification task was used because it is characterized by a high degree of similarity in unspecific task requirements (i.e., forced choice selection from five different “elementary” categories). The observed activation of the right frontal and temporal cortex during the emotional task as compared to the phonological control condition provides support for the hypothesis of a specific right hemisphere contribution to processing of intonational information derived from clinical investigations (Baum and Pell, 1999; Borod et al., 2001, 2002; Charbonneau et al., 2003; Darby, 1993; Heilman et al., 1984; Kucharska-Pietura et al., 2003; Pell and Baum, 1997; Starkstein et al., 1994) and prior functional imaging studies (Buchanan et al., 2000; George et al., 1996; Meyer et al., 2002; Mitchell et al., 2003; Pihan et al., 1997, 2000; Wildgruber et al., 2002, 2004).

The activation cluster within the right temporal cortex is localized within the superior temporal sulcus (BA 22/42). A comparison of the signal time curves during the emotional and the phonetic task reveals activation of this region during both tasks in all subjects participating in the study (Fig. 5). This pattern of hemodynamic responses during the emotional identification task is in good accordance with the assumption that the right auditory association cortex predominantly contributes to the analysis of acoustic features that are important to distinguish the emotional state of the speaker as indicated by the majority of the available lesion studies (Borod et al., 2002; Breitenstein et al., 1998; Darby, 1993; Heilman et al., 1984; Starkstein et al., 1994) as well as previous functional imaging studies (Buchanan et al., 2000; Kotz et al., 2003; Mitchell et al., 2003; Wildgruber et al., 2004). Rightward lateralization within the posterior temporal region, however, does not seem to be associated to the emotional significance of the acoustic signal. Findings from several recent studies rather indicate the right temporal region to be specifically involved in the analysis of slow acoustic variations characterizing suprasegmental aspects of the speech signal, whereas the left temporal region has been linked to analysis of rapid acoustic changes required for the differentiation of speech sounds at phoneme level (Ackermann et al., 2001; Ivry and Robertson, 1998; Johnsrude et al., 2000; Meyer

et al., 2002; Poeppel et al., 2004; Van Lancker and Sidtis, 1992; Zatorre, 2001).

Activation within the right dorsolateral frontal cortex (BA 9/45/46) has been observed during discrimination of pure tonal patterns in previous imaging studies (Zatorre, 2001; Zatorre et al., 1992, 1994). Since processing of emotional intonation relies on extraction, storage, and comparison of fundamental frequency variations, recruitment of a pitch working memory system must be considered essential for successful task performance. Responses within the right dorsolateral frontal region observed in the current experiment and in prior imaging studies during processing of speech prosody (Buchanan et al., 2000; George et al., 1996; Kotz et al., 2003; Mitchell et al., 2003; Wildgruber et al., 2002, 2004), therefore, presumably reflect engagement of this system. In line with this assumption, the clinical observations of impaired prosody comprehension after damage to the right frontal cortex (Adolphs et al., 2002; Baum and Pell, 1999; Borod et al., 2002; Breitenstein et al., 1998) might be linked to a deficit of the pitch working memory system.

Responses within the right orbitobasal cortex (BA 47) were linked to comprehension of emotional prosody in the current study. Activation of this region has been reported during perception of emotional intonation (George et al., 1996; Wildgruber et al., 2002, 2004), emotional facial expressions (Blair et al., 1999; Nakamura et al., 1999), and affective gustatory judgements (Small et al., 2001) in prior functional imaging studies. Moreover, patients suffering from unilateral circumscribed lesions to this area showed impaired identification of emotional face and voice expressions whereas performance in nonemotional control tasks (i.e., discrimination of unfamiliar voices and recognition of environmental sounds) was unimpaired (Hornak et al., 1996, 2003; Rolls, 1999). These observations—in line with the results of the current study—support the presumption that the orbitofrontal region contributes to the analysis of emotional information conveyed by different communicational channels.

*Influence of specific emotions*

No significant differences of brain activation depending on emotional valence or specific emotional categories expressed by speech intonation could be observed. Therefore, the results of the current study in line with prior functional imaging examinations (Buchanan et al., 2000; Kotz et al., 2003; Mitchell et al., 2003; Wildgruber et al., 2002) and recent lesion studies (Baum and Pell, 1999; Borod et al., 2002; Kucharska-Pietura et al., 2003; Pell, 1998) do not support the hypothesis of valence-specific lateralization effects during processing of emotional intonation. Considering hemodynamic responses bound to specific emotional categories, a selective contribution of the amygdala to auditory recognition of fear has been assumed on the basis of lesion data (Scott et al., 1997) and prior PET studies (Morris et al., 1999; Phillips et al., 1998). Furthermore, a specific association of responses within the anterior insula and the basal ganglia during perception of vocal expressions of disgust has been predicted based on clinical findings (Pell and Leonard, 2003) and functional imaging experiments during processing of facial expressions (Phan et al., 2002; Sprengelmeyer et al., 1998; Wicker et al., 2003). The present study, however, failed to reveal fear-related amygdala activity as well as disgust-specific activation of the insula or basal ganglia. Possibly, the absence of statistically significant differences of cerebral responses with respect to

specific emotional categories might be related to low statistical power due to the reduced sample size for the analysis of subdivisions of the acquired images (as compared to task vs. baseline and task 1 vs. task 2 contrasts). Moreover, responses of the amygdalae have been observed to depend on passive or implicit processing of emotional signals, whereas explicit judgments of emotional expressions, as required in the current study, have been shown to result in deactivation concomitant with suppression of psychophysiological responses (Adolphs, 2002; Morris et al., 1999). Activation of other brain regions associated to specific emotional categories might also be linked to implicit processing and induction of the respective emotional states. The fine nuances of emotional tone considered in the present study, however, were not intended to evoke strong emotional reactions and induction of specific emotions has not been reported by the participants.

## Conclusion

The findings of the present study support the hypothesis of an important contribution of right-sided temporal and frontal regions to the processing of emotional prosody independent of specific emotional categories. The observed rightward lateralization at the level of the posterior temporal cortex (BA 22/42) might be bound to extraction of specific acoustic cues from complex speech signals (i.e., suprasegmental features such as pitch contours and rhythmic structures), whereas engagement of the orbitofrontal cortex (BA 47) presumably reflects a contribution to evaluation of emotional associations.

## Acknowledgments

This study was supported by the Junior Science Program of the Heidelberg Academy of Sciences and Humanities and the German Research Foundation (DFG WI 2101).

## References

- Akermann, H., Hertrich, I., Ziegler, W., 1993. Prosodische störungen bei neurologischen erkrankungen-eine literaturübersicht. *Fortschr. Neurol. Psychiatr.* 61, 241–253.
- Akermann, H., Riecker, A., Mathiak, K., Erb, M., Grodd, W., Wildgruber, D., 2001. Rate-dependent activation of a prefrontal–insular–cerebellar network during passive listening to trains of click stimuli: an fMRI study. *NeuroReport* 12, 4087–4092.
- Adolphs, R., 2002. Neural systems for recognizing emotion. *Curr. Opin. Neurobiol.* 12, 169–177.
- Adolphs, R., Damasio, H., Tranel, D., 2002. Neural systems for recognition of emotional prosody: a 3-D lesion study. *Emotion* 2, 23–51.
- Bachorowski, J.O., Owren, M.J., 2003. Sounds of emotion: production and perception of affect-related vocal acoustics. *Ann. N. Y. Acad. Sci.* 1000, 244–265.
- Banse, R., Scherer, K.R., 1996. Acoustic profiles in vocal emotion expression. *J. Pers. Soc. Psychol.* 70, 614–636.
- Baum, S.R., Pell, M.D., 1999. The neural basis of prosody: insights from lesion studies and neuroimaging. *Aphasiology* 13, 581–608.
- Baumgart, F., Kaulisch, T., Tempelmann, C., Gaschler-Merkel, B., Tegeler, C., Schindler, F., Stiller, D., Schaich, H., 1998. Electrodynamical headphones and woofers for application in magnetic resonance imaging scanners. *Med. Phys.* 25, 2068–2070.
- Blair, R.J.R., Morris, J.S., Frith, C.D., Perret, D.I., Dolan, R.J., 1999. Dissociable neural responses to facial expressions of sadness and anger. *Brain* 122, 883–893.
- Borod, J.C., Zgaljardic, D., Tabert, M.H., Koff, E., 2001. Asymmetries of emotional perception and expression in normal adults. In: Gainotti, G. (Ed.), *Handbook of Neuropsychology*, second ed., vol. 5. Elsevier Science, pp. 181–205.
- Borod, J.C., Bloom, R.L., Brickman, A.M., Nakhutina, L., Curko, E.A., 2002. Emotional processing deficits in individuals with unilateral brain damage. *Appl. Neuropsychol.* 9, 23–36.
- Breitenstein, C., Daum, I., Ackermann, H., 1998. Emotional processing following cortical and subcortical brain damage: contribution of the fronto-striatal circuitry. *Behav. Neurol.* 11, 29–42.
- Breitenstein, C., Van Lancker, D., Daum, I., Waters, C., 2001. Impaired perception of vocal emotions in Parkinson's disease: influence of speech time processing and executive functioning. *Brain Cogn.* 45, 277–314.
- Buchanan, T.W., Lutz, K., Mirzazade, S., Specht, K., Shah, N.J., Zilles, K., Jäncke, L., 2000. Recognition of emotional prosody and verbal components of spoken language: an fMRI study. *Cogn. Brain Res.* 9, 227–238.
- Calder, A.J., Keane, J., Manes, F., Antoun, N., Young, A.W., 2000. Impaired recognition and experience of disgust following brain injury. *Nat. Neurosci.* 3, 1077–1078.
- Cancelliere, A.E., Kertesz, A., 1990. Lesion localization in acquired deficits of emotional expression and comprehension. *Brain Cogn.* 13, 133–147.
- Canli, T., Desmond, J.E., Zhao, Z., Glover, G., Gabrieli, J.D., 1998. Hemispheric asymmetry for emotional stimuli detected with fMRI. *NeuroReport* 9, 3233–3239.
- Charbonneau, S., Scherzer, B.P., Aspirot, D., Cohen, H., 2003. Perception and production of facial and prosodic emotions by chronic CVA patients. *Neuropsychologia* 41, 605–613.
- Collins, D.L., Neelin, P., Peters, T.M., Evans, A.C., 1994. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *J. Comput. Assist. Tomogr.* 18, 192–205.
- Cutler, A., Dahan, D., Donselaar, W., 1997. Prosody in the comprehension of spoken language: a literature review. *Lang. Speech* 40, 141–201.
- Darby, D.G., 1993. Sensory aprosodia: a clinical clue to lesions of the inferior division of the right middle cerebral artery? *Neurology* 43, 567–572.
- Davidson, R.J., Tomarken, A.J., 1989. Laterality and emotion: an electrophysiological approach. In: Boller, F., Grafman, J. (Eds.), *Handbook of Neuropsychology*. Elsevier, Amsterdam, pp. 419–441.
- Davidson, R.J., Abercrombie, H., Nitschke, J.B., Putnam, K., 1999. Regional brain function, emotion and disorders of emotion. *Curr. Opin. Neurobiol.* 9, 228–234.
- George, M.S., Parekh, P.I., Rosinsky, N., Ketter, T.A., Kimbrell, T.A., Heilman, K.M., Herscovitch, P., Post, R.M., 1996. Understanding emotional prosody activates right hemisphere regions. *Arch. Neurol.* 53, 665–670.
- Heilman, K.M., Bowers, D., Speedie, L., Coslett, H.B., 1984. Comprehension of affective and nonaffective prosody. *Neurology* 34, 917–921.
- Hornak, J., Rolls, E.T., Wade, D., 1996. Face and voice expression identification in patients with emotional and behavioral changes following ventral frontal lobe damage. *Neuropsychologia* 34, 247–261.
- Hornak, J., Bramham, J., Rolls, E.T., Morris, R.G., O'Doherty, J., Bullock, P.R., Polkey, C.E., 2003. Changes in emotion after circumscribed surgical lesions of the orbitofrontal and cingulate cortices. *Brain* 126, 1691–1712.
- Imaizumi, S., Mori, K., Kiritani, S., Kawashima, R., Sugiura, M., Fukuda, H., Itoh, K., Kato, T., Nakamura, A., Hatano, K., Kojima, S., Nakamura, K., 1997. Vocal identification of speaker and emotion activates different brain regions. *NeuroReport* 8, 2809–2812.
- Ivry, R.B., Robertson, L.C., 1998. *The Two Sides of Perception*. MIT Press, Cambridge, MA.
- Johnsrude, I.S., Penhune, V.B., Zatorre, R.J., 2000. Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain* 123, 155–163.

- Kesler-West, M.L., Andersen, A.H., Smith, C.D., Avison, M.J., Davis, C.E., Kryscio, R.J., Blonder, L.X., 2001. Neural substrates of facial emotion processing using fMRI. *Cogn. Brain Res.* 11, 213–226.
- Klose, U., Erb, M., Wildgruber, D., Müller, E., Grodd, W., 1999. Improvement of the acquisition of a large amount of MR images on a conventional whole body system. *Magn. Reson. Imaging* 17, 471–474.
- Kotz, S.A., Meyer, M., Alter, K., Besson, M., von Cramon, D.Y., Friederici, A.D., 2003. On the lateralization of emotional prosody: an event-related functional MR investigation. *Brain Lang.* 86, 366–376.
- Kucharska-Pietura, K., Phillips, M.L., Germand, W., David, A.S., 2003. Perception of emotions from faces and voices following unilateral brain damage. *Neuropsychologia* 41, 1082–1090.
- Meyer, M., Alter, K., Friederici, A.D., Lohmann, G., von Cramon, D.Y., 2002. fMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. *Hum. Brain Mapp.* 17, 73–88.
- Mitchell, R.L.C., Elliot, R., Barry, M., Cruttenden, A., Woodruff, P.W.R., 2003. The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. *Neuropsychologia* 41, 1410–1421.
- Morris, J.S., Scott, S.K., Dolan, R.J., 1999. Saying it with feeling: neural responses to emotional vocalization. *Neuropsychologia* 37, 1155–1163.
- Murphy, F.C., Nimmo-Smith, I., Lawrence, A.D., 2003. Functional neuroanatomy of emotions: a meta-analysis. *Cogn. Affect Behav. Neurosci.* 3, 207–233.
- Nakamura, K., Kawashima, R., Ito, K., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Nagumo, S., Kubota, K., Fukuda, H., Kojima, S., 1999. Activation of the right inferior frontal cortex during assessment of facial emotion. *J. Neurophysiol.* 82, 1610–1614.
- Newman, S.D., Twieg, D.B., Carpenter, P.A., 2001. Baseline conditions and subtractive logic in neuroimaging. *Hum. Brain Mapp.* 14, 228–235.
- Oldfield, R.C., 1971. Assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.
- Paulesu, E., Frith, C.D., Frackowiak, R.S.J., 1993. The neural correlates of the verbal component of working memory. *Nature* 362, 342–345.
- Pell, M.D., 1998. Recognition of prosody following unilateral brain lesions: influence of functional and structural attributes of prosodic contours. *Neuropsychologia* 36, 701–715.
- Pell, M.D., Baum, S.R., 1997. Unilateral brain damage, prosodic comprehension deficits, and the acoustic cues to prosody. *Brain Lang.* 57, 195–214.
- Pell, M.D., Leonard, C.L., 2003. Processing emotional tone from speech in Parkinson's disease: a role for the basal ganglia. *Cogn. Affect. Behav. Neurosci.* 3, 275–288.
- Phan, K.L., Wager, T., Taylor, S.F., Liberzon, I., 2002. Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage* 16, 331–348.
- Phillips, M.L., Young, A.W., Scott, S.K., Calder, A.J., Andrew, C., Giampietro, V., Williams, S.C.R., Bullmore, E.T., Brammer, M., Gray, J.A., 1998. Neural responses to facial and vocal expressions of fear and disgust. *Proc. R. Soc. London* 265, 1809–1817.
- Pihan, H., Altenmüller, E., Ackermann, H., 1997. The cortical processing of perceived emotion: a DC-potential study on affective speech 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.
- Poeppl, D., 1996. A critical review of PET studies of phonological processing. *Brain Lang.* 55, 317–351.
- Poeppl, D., Guillemin, A., Thompson, J., Fritz, J., Bavelier, D., Braun, A., 2004. Auditory lexical decision, categorical perception, and FM direction discrimination differentially engage left and right auditory cortex. *Neuropsychologia* 42, 183–200.
- Price, D.L., De Wilde, J.P., Papadaki, A.M., Curran, J.S., Kitney, R.I., 2001. Investigation of acoustic noise on 15 MRI scanners from 0.2 T to 3 T. *J. Magn. Reson. Imaging* 13, 288–293.
- Rolls, E.T., 1999. The functions of the orbito-frontal cortex. *Neurocase* 5, 301–312.
- Ross, E.D., 1981. The aprosodias: functional-anatomic organization of the affective components of language in the right hemisphere. *Arch. Neurol.* 38, 561–569.
- Scott, S.K., Young, A.W., Calder, A.J., Hellawell, D.J., Aggleton, J.P., Johnson, M., 1997. Impaired auditory recognition of fear and anger following bilateral amygdala lesions. *Nature* 385, 254–257.
- Shallice, T., Vellar, G., 1990. The impairment of auditory-verbal short-term storage. In: Vellar, G., Shallice, T. (Eds.), *Neuropsychological Impairments of Short-term Memory*. Cambridge Univ. Press, Cambridge, pp. 11–53.
- Sidtis, J.J., Van-Lancker-Sidtis, D., 2003. A neurobehavioral approach to dysprosody. *Semin. Speech Lang.* 24, 93–105.
- Small, D.M., Zatorre, R.J., Dagher, A., Evans, A.C., Jones-Gotman, M., 2001. Changes in brain activity related to eating chocolate: from pleasure to aversion. *Brain* 124, 1720–1733.
- Sprengelmeyer, R., Rausch, M., Eysel, U.T., Przuntek, H., 1998. Neural structures associated with recognition of facial expressions of basic emotions. *Proc. R. Soc. Lond., B Biol. Sci.* 265, 1927–1931.
- Stark, C.E., Squire, L.R., 2001. When zero is not zero: the problem of ambiguous baseline conditions in fMRI. *Proc. Natl. Acad. Sci. U. S. A.* 98, 12760–12766.
- Starkstein, S.E., Federoff, J.P., Price, T.R., Leiguarda, R.C., Robinson, R.G., 1994. Neuropsychological and neuroradiologic correlates of emotional prosody comprehension. *Neurology* 44, 515–522.
- Van Lancker, D., Sidtis, J.J., 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.
- Wicker, B., Keysers, C., Plailly, J., Royet, J.P., Gallese, V., Rizzolatti, G., 2003. Both of us disgusted in my insula: the common neural basis of seeing and feeling disgust. *Neuron* 40, 655–664.
- Wildgruber, D., Kischka, U., Ackermann, H., Klose, U., Grodd, W., 1999. Dynamic pattern of brain activation during sequencing of word strings evaluated by fMRI. *Cogn. Brain Res.* 7, 285–294.
- Wildgruber, D., Pihan, H., Ackermann, H., Erb, M., Grodd, W., 2002. Dynamic brain activation during processing of emotional intonation: influence of acoustic parameters, emotional valence and sex. *NeuroImage* 15, 856–869.
- Wildgruber, D., Hertrich, I., Riecker, A., Erb, M., Anders, S., Grodd, W., Ackermann, H., 2004. Distinct frontal regions subserve evaluation of linguistic and emotional aspects of speech intonation. *Cereb. Cortex* 14, 1384–1389.
- Zatorre, R.J., 2001. Neural specializations for tonal processing. *Ann. N. Y. Acad. Sci.* 930, 193–210.
- 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., Evans, A.C., Meyer, E., 1994. Neural mechanisms underlying melodic perception and memory for pitch. *J. Neurosci.* 14, 1908–1919.