Specific Brain Networks during Explicit and Implicit Decoding of Emotional Prosody

Sascha Frühholz[1,2], Leonardo Ceravolo[1,2] and Didier Grandjean[1,2]

[1]Neuroscience of Emotion and Affective Dynamics Lab, Department of Psychology and [2]Swiss Center for Affective Sciences, University of Geneva, CH-1205 Genèva, Switzerland

Address correspondence to Dr Sascha Frühholz, Neuroscience of Emotion and Affective Dynamics Lab, Department of Psychology, University of Geneva, 40bd du Pont d'Arve, CH-1205 Genève, Switzerland. Email: sascha.fruehholz@unige.ch.

To better define the underlying brain network for the decoding of emotional prosody, we recorded high-resolution brain scans during an implicit and explicit decoding task of angry and neutral prosody. Several subregions in the right superior temporal gyrus (STG) and bilateral in the inferior frontal gyrus (IFG) were sensitive to emotional prosody. Implicit processing of emotional prosody engaged regions in the posterior superior temporal gyrus (pSTG) and bilateral IFG subregions, whereas explicit processing relied more on mid STG, left IFG, amygdala, and subgenual anterior cingulate cortex. Furthermore, whereas some bilateral pSTG regions and the amygdala showed general sensitivity to prosody-specific acoustical features during implicit processing, activity in inferior frontal brain regions was insensitive to these features. Together, the data suggest a differentiated STG, IFG, and subcortical network of brain regions, which varies with the levels of processing and shows a higher specificity during explicit decoding of emotional prosody.

Keywords: acoustic stimulus features, emotional prosody, fMRI, frontotemporal network, implicit/explicit processing

Introduction

The human brain incorporates a frontotemporal network of regions that decode affective cues and infer emotional states from suprasegmental vocal modulations, referred to as emotional prosody (Banse and Scherer 1996; Grandjean et al. 2006). Specifically, regions along the superior temporal gyrus (STG) and the superior temporal sulcus (STS) as well as in the orbitofrontal cortex (OFC) and the ventrolateral prefrontal cortex have been found to be involved in the sensory, emotional, and evaluative decoding of emotional prosody (Schirmer and Kotz 2006; Wildgruber et al. 2009). In addition to this frontotemporal network, the amygdala (Grandjean et al. 2005; Sander et al. 2005; Fecteau et al. 2007; Bach et al. 2008) and subcortical regions, such as the thalamus (Wildgruber et al. 2004) and the basal ganglia (Kotz et al. 2003; Bach et al. 2008; Grandjean et al. 2008), have been found to be sensitive to emotional prosody. However, evidence for sensitivity of the latter regions to emotional prosody is inconsistent (Calder et al. 2004; Mitchell and Boucas 2009), especially for the amygdala. Some studies report activations of the amygdala (Scott et al. 1997; Sprengelmeyer et al. 1999; Wiethoff et al. 2009; Leitman et al. 2010), whereas the results from other studies do not support its involvement in the processing of emotional prosody (Adolphs and Tranel 1999; Anderson and Phelps 2002; Belin et al. 2008).

Although these recent studies provided evidence for the decoding of emotional information from vocal cues in the frontotemporal and subcortical regions, precise spatial information, especially about the involvement of different temporal and frontal subregions, is still lacking. The location of peak activation can vary substantially and usually extends broadly across different regions. Within the temporal cortex, studies report peak activations for emotional compared with neutral prosody that range from the bilateral posterior superior temporal gyrus (pSTG; Mitchell et al. 2003; Sander et al. 2005; Ethofer, Anders, Erb, et al. 2006; Beaucousin et al. 2007; Ethofer, Kreifelts, et al. 2009) and the bilateral midsuperior temporal gyrus (mSTG; Grandjean et al. 2005; Sander et al. 2005; Bach et al. 2008; Ethofer, Kreifelts, et al. 2009; Leitman et al. 2010) to the right anterior superior temporal gyrus (aSTG; Bach et al. 2008). Although this variation in peak location follows the distribution of human voice-sensitive areas in the superior temporal cortex (Belin et al. 2000), it might suggest different functional roles of these subregions in decoding affective cues from voices (Schirmer and Kotz 2006). For example, adjacent regions in the right pSTG and mSTG are sensitive to different levels of processing. That means, some studies found a general sensitivity in the pSTG or mSTG independent of the attentional focus (directed to or away from emotional prosody), while explicitly focusing attention towards emotional prosody revealed activations in adjacent regions of the STG (Grandjean et al. 2005; Sander et al. 2005). Apart from this variation in peak location, these activations usually extend broadly in the STG or STS and sometimes extend to or comprise regions in the primary and secondary auditory cortex (Wildgruber et al. 2004; Sander et al. 2005; Leitman et al. 2010). The latter are more involved in decoding basic sensory acoustic features from affective utterances rather than in decoding higher level affective cues.

Similar to the variation and extent of activations in the superior temporal cortex, activations in the bilateral inferior frontal gyrus (IFG) that cover different frontal subregions have been frequently reported. Activations were found in the bilateral pars opercularis of the IFG (Brodmann area [BA] 44/45; Buchanan et al. 2000; Mitchell et al. 2003; Schirmer et al. 2004), but these activations often extend to the pars orbitalis of the IFG (BA 47; Fecteau et al. 2005; Bach et al. 2008; Ethofer, Kreifelts, et al. 2009) and to the middle frontal gyrus (BA 9/46; Wildgruber et al. 2004; Ethofer, Anders, Erb, et al. 2006; Leitman et al. 2010). These different inferior frontal subregions are again assumed to serve different functional roles. Left BA 44 is generally involved in linguistic and semantic processes (Kotz et al. 2003; Schirmer et al. 2004) and is functionally distinct from BA 47, which is specifically sensitive to the emotional tone of a voice (Fecteau et al. 2005). Furthermore, right BA 45 has been proposed to be involved in the cognitive evaluation of emotional stimuli, whereas BA 47 decodes the reward value of emotional stimuli when activations also comprise regions in the OFC (Schirmer and Kotz 2006). Right BA 47 might also assess

and mirror the emotional quality of auditory cues as was shown for emotional faces (Nakamura et al. 1999; Lee et al. 2006).

Therefore, given the variation of functional activations and the different functional properties of STG and IFG subregions, a more precise anatomical description of these activations and their functional roles is needed. In the present study, we used high-resolution functional magnetic resonance imaging (fMRI) scans to better separate and determine functional activations during the decoding of emotional prosody. Furthermore, we were interested in the dependency of these functional activations on the level of processing of emotional utterances. Finally, we were also interested in the dependency of functional activations on those acoustical stimulus features that support the formation of emotional representations from acoustic cues, such as the fundamental frequency (F0) perceived as pitch and the energy/intensity (I) perceived as intensity of acoustic stimuli. Both features show a specific pattern, depending on the emotional valence (Banse and Scherer 1996; Grandjean et al. 2006).

Concerning the level of processing, this level can range from explicit judgments on emotional prosody to passive listening to stimuli and finally to rather implicit levels of processing when attention is directed toward other stimulus features (Bach et al. 2008) or spatial locations (Grandjean et al. 2005; Sander et al. 2005). This level of processing refers to what has formerly been described as the appraisal level during processing of emotional prosody (Bach et al. 2008) and concerns the attentional focus, which can be directed towards (explicit decoding) or away from emotional prosody (implicit decoding). As mentioned earlier, activation is shifted to adjacent regions of the right midposterior superior temporal gyrus (m-pSTG) when the task requires an explicit rather than a task-independent decoding of emotional prosody, which induces activity in m-pSTG (Grandjean et al. 2005; Sander et al. 2005). This might also suggest task-dependent differences in functional activations in the wider frontotemporal network. Explicit decoding, compared with implicit decoding, induces activity in the bilateral pSTG (Mitchell et al. 2003; Sander et al. 2005; Wildgruber et al. 2005; Ethofer, Kreifelts, et al. 2009), bilateral IFG (BA 45/47) (Buchanan et al. 2000; Wildgruber et al. 2005; Bach et al. 2008; Ethofer, Kreifelts, et al. 2009), and OFC (Wildgruber et al. 2004; Sander et al. 2005), whereas implicit processing is accompanied by activity in the bilateral pSTG (Sander et al. 2005; Beaucousin et al. 2007; Bach et al. 2008), right aSTG (Mitchell et al. 2003), bilateral amygdala (Sander et al. 2005; Bach et al. 2008), and left IFG (Buchanan et al. 2000; Wildgruber et al. 2004; Bach et al. 2008; Ethofer, Kreifelts, et al. 2009), the latter probably due to the focus on linguistic and semantic stimulus features while ignoring the prosodic stimulus feature.

Activity in the frontotemporal regions, in addition to task-dependent levels of processing, might also depend on emotion-specific acoustic stimulus features. Although some studies did not find a dependency on functional activations for basic stimulus features (Wildgruber et al. 2002), others report a covariation of bilateral mSTG activity with the perceived emotional intensity of prosodic stimuli based on F0 variability (Ethofer, Anders, Wiethoff, et al. 2006) or with mean F0 in the left mSTG and with mean energy in the right mSTG (Wiethoff et al. 2008). Increasing emotion-specific acoustic cues, such as high-frequency cues of angry stimuli or F0 variability for fearful and happy stimuli, increases activity in the bilateral pSTG with

a concurrent signal decrease in the bilateral IFG (Leitman et al. 2010). These data might suggest a general sensitivity of STG regions to emotion-specific acoustic cues as the basis of auditory emotional representations. However, assuming that more pSTG regions decode sensory cues and more aSTG regions decode emotional cues (Schirmer and Kotz 2006), these data might suggest a distinction in STG regions in their sensitivity to acoustic stimulus features. Furthermore, the amygdala and subcortical structures might additionally be sensitive to emotion-specific acoustic features. Among other brain regions, the amygdala might decode the emotional value from acoustic stimuli (Sander et al. 2005; Fecteau et al. 2007; Bach et al. 2008; Ethofer, Kreifelts, et al. 2009; Wiethoff et al. 2009), which might be based on emotion-specific acoustic cues (Leitman et al. 2010). Similarly, the basal ganglia are sometimes involved in emotional prosody decoding (Kotz et al. 2003; Bach et al. 2008; Grandjean et al. 2008) on the basis of analysis and integration of temporal patterns (Kotz and Schwartze 2010). However, this sensitivity of different STG regions, of the amygdala or of the basal ganglia to emotion-specific acoustic cues, is still unknown.

Therefore, the present study had 3 general aims. First, we used high-resolution fMRI scans to describe the frontotemporal and subcortical network involved in emotional prosody decoding on a finer spatial scale. We predicted that the cluster of regions in the STG and IFG would consist of several subregions that might subserve different functional roles. Second, if decoding of emotional prosody in the STG and IFG is accomplished by different subregions, we hypothesized that activity in these regions might depend on the level of processing (Grandjean et al. 2005; Sander et al. 2005). We used an explicit and an implicit task during the processing of the same prosodic stimuli (Fig. 1A). The explicit task involved a discrimination of prosodic stimuli (anger vs. neutral) and should elicit activity in more aSTG and especially in frontal brain regions (Sander et al. 2005; Bach et al. 2008; Ethofer, Kreifelts, et al. 2009). The implicit task involved a gender discrimination of the speaker's voice. Gender processing of voices, or identity processing more general, usually elicits activity in more aSTG regions (Belin and Zatorre 2003; Formisano et al. 2008), which is independent of the emotional tone of a voice. However, beyond these simple task effects, we specifically expected that both levels of processing will differentially influence the decoding of emotional compared with neutral prosody. We specifically expected that a more explicit decoding of emotional compared with neutral prosody might reveal activation more anteriorly in STG regions compared with implicit processing (Grandjean et al. 2005). Since different emotional prosodies comprise a different pattern of basic acoustic features, we finally explored the influence of the mean and variation of F0 and the energy of prosodic stimuli on functional activations during both levels of processing, and we predicted a stronger influence on brain regions that decode sensory stimulus features compared with higher level processing regions.

Materials and Methods

Participants
Seventeen healthy participants recruited from Geneva University took part in the experiment (3 males; mean age 25.52 years, standard deviation [SD] = 5.08, age range 20-38 years). All participants were
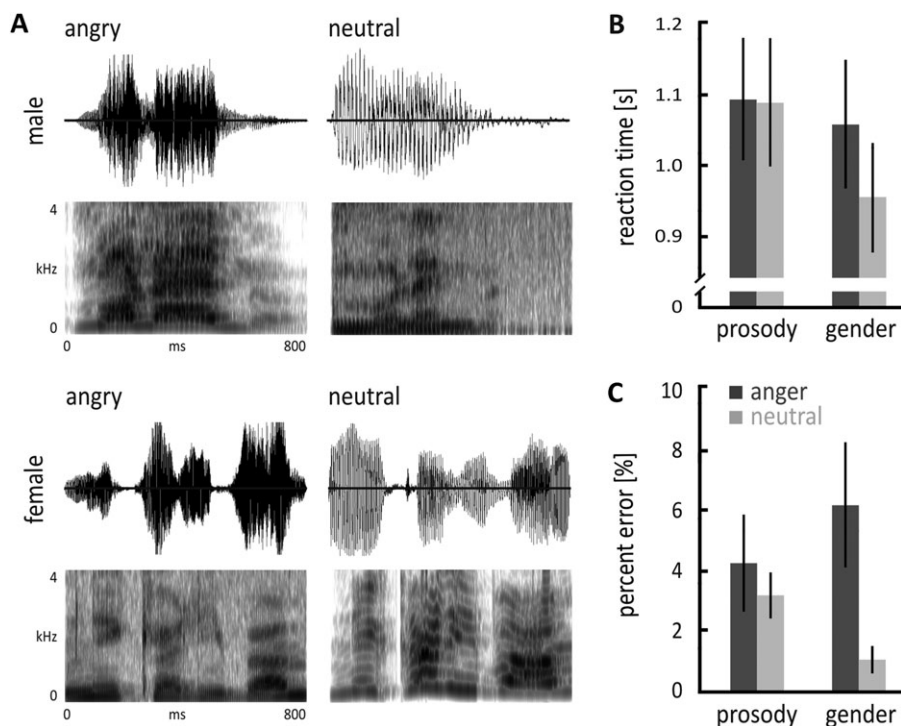
Figure 1. Participants heard 4 different speech like but semantically meaningless words spoken in an angry or neutral tone by 2 male and 2 female actors. During 2 blocks, participants had to make prosody discriminations ("angry" or "neutral") on the voices, and during another 2 blocks, they had to decide about the gender of the voice ("male" or "female"). (*A*) The upper panel shows amplitude waveforms and spectrograms for the word /molen/ spoken by a male voice in an angry (left) and neutral tone (right). The lower panel shows the word /nikalibam/ spoken in an angry and neutral tone by a female voice. (*B*) Participants generally made slower response during the prosody (angry voices [SEM 86], neutral voices [SEM 90]) compared with the gender discrimination task (angry voices [SEM 90], neutral voices [SEM 77]; $F_{1,16} = 7.072$, $P = 0.017$) and for angry compared with neutral voices ($F_{1,16} = 6.558$, $P = 0.021$). Error bars denote the SEM. (*C*) Performance accuracy did not differ between both tasks ($F_{1,16} < 1$) and between emotional stimuli ($F_{1,16} = 2.700$, $P = 0.120$). A task × emotion interaction ($F_{1,16} = 6.561$, $P = 0.021$), however, indicated that participants specifically made more errors during gender discriminations of angry compared with neutral voices (angry voices 6.43% [SEM 2.15], neutral voices 1.10% [SEM 0.46]; $t_{16} = 2.365$, $P = 0.031$) compared with prosody discriminations on the same stimuli (angry voices 4.41% [SEM 1.65], neutral voices 3.31% [SEM 0.78]).

native French speakers, were right handed, had normal or corrected-to-normal vision, and had normal hearing abilities. No participant presented a neurologic or psychiatric history. All participants gave informed and written consent for their participation in accordance with ethical and data security guidelines of the University of Geneva. The study was approved by the local ethics committee.

Stimulus Material and Trial Sequence

The stimulus material consisted of 4 speech like but semantically meaningless words ("molen," "belam," "nikalibam," and "kudsemina") extracted from the Geneva Multimodal Emotion Portrayal database (Bänziger and Scherer 2010). These words were spoken in either a neutral or an angry tone by 2 male and 2 female speakers, resulting in 32 different stimuli (see Fig. 1*A*). Auditory stimuli had a mean duration of 690 ms and were equated for mean sound pressure level. A preevaluation of the stimuli by 16 participants (5 males; mean age 25 years, SD = 6.15, age range 20–39 years) revealed that both neutral and angry stimuli were significantly rated as neutral ($F_{1,15} = 188.464$, $P = 8.835 \times 10^{-32}$) and angry ($F_{4,56} = 163.692$, $P = 3.380 \times 10^{-30}$), respectively. Angry voices were rated as more arousing compared with neutral stimuli ($F_{1,14} = 88.371$, $P = 1.998 \times 10^{-7}$).

During scanning, auditory stimuli were presented binaurally with magnetic resonance imaging compatible headphones (MR Confon) at a sound pressure level of approximately 70 dB. Auditory stimuli were preceded by a visual fixation cross (1 × 1°) for 1 ± 0.5 s. The fixation cross-remained on the screen for as long as 2 s after the auditory stimulus to indicate a time window during which participants were stressed to respond. Auditory stimuli were presented between functional volume acquisitions and had an onset of 4 ± 0.75 s prior to the onset of the volume acquisition that followed (see below).

The same stimuli were presented during 2 blocks of explicit prosody discriminations on the stimuli (neutral or angry; right index and middle finger) and during another 2 blocks, where participants made gender discriminations (male or female) on the voices. For the latter, we assumed that though participants focus on the gender of the voice, the emotional tone of the voice is still decoded on an implicit level of processing. This task might involve functional activations, which are related to the explicit processing of gender information of the voice. However, as for the explicit task, we also computed contrasts between angry and neutral voices for this implicit task, which keeps the gender of the voice counterbalanced for this comparison. This should eliminate any activation primarily related to processing of the gender of a voice. Each of the experimental blocks contained 38 trials, including 6 silent events with no auditory stimulation. Task blocks alternated across the experiment, and block order and response buttons were counter-balanced across participants.

To localize human voice-sensitive regions in the bilateral superior temporal cortex, we used 8-s sound clips taken from an existing database (see http://vnl.psy.gla.ac.uk/ and Belin et al. 2000). These sounds clips contained 20 sequences of human voices and 20 sequences of animal or environmental sounds. Each sound clip was presented once with a fixation cross on the screen and a 4-s gap between each clip. The scanning sequence also contained twenty 8-s silent events, and participants had to passively listen to the stimuli (see Supplementary Material for a full description of the voice localizer scan).

Image Acquisition

For the main experiment, we obtained high-resolution imaging data on a 3 T Siemens Trio System (Siemens, Erlangen, Germany) by using a T2*-weighted gradient echo planar imaging sequence. Twenty-five

axial slices were aligned to the superior temporal sulcus along the anterior–posterior orientation (thickness/gap = 2/0.4 mm, field of view (FoV) = 192 mm, in-plane 1.5 × 1.5 mm). We used a sparse temporal acquisition protocol with time repetition (TR) = 10 s, which consisted of 1.75 s for volume acquisition and a silent gap of 8.25 s. For the voice localizer, we used a continuous whole-head acquisition of 36 slices (thickness/gap = 3.2/0.64 mm, FoV = 205 mm, in-plane 3.2 × 3.2 mm) aligned to the anterior to posterior commissure plane with TR/time echo (TE) = 2.1/0.03 s. Finally, a high-resolution magnetization prepared rapid acquisition gradient echo T1-weighted sequence (192 contiguous 1 mm slices, TR/TE/time to inversion = 1.9 s/2.27 ms/900 ms, FoV = 296 mm, in-plane 1 × 1 mm) was obtained in sagittal orientation to obtain structural brain images from each participant.

Image Analysis
We used the statistical parametric mapping software SPM (version 8; Welcome Department of Cognitive Neurology, London, UK) for preprocessing and statistical analysis of functional images. Functional images were realigned and coregistered to the anatomical image. A segmentation of the anatomical image revealed warping parameters that were used to normalize the functional images to the Montreal Neurological Institute (MNI) stereotactic template brain. During normalization, functional images from the voice localizer scan were resampled to 1.5 × 1.5 × 2 mm voxel size. Normalized images were spatially smoothed with a nonisotropic Gaussian kernel of full-width at half-maximum 3 × 3 × 4 mm.

We used a general linear model for the first-level statistical analyses, including boxcar functions defined by the onset and duration of the auditory stimuli. These boxcar functions were convolved with a canonical hemodynamic response function. Separate regressors were created for each experimental condition, and the general linear model for the main experiment also included one additional repressor containing all erroneous and missed trials. Six motion correction parameters were finally included as regressors of no interest to minimize false-positive activations that were due to task-correlated motion. Linear contrasts for the experimental conditions for each participant were taken to a second-level random effects group analysis of variance.

For the main experiment, we set up a 2 × 2 factorial design including the factors "task" and "emotion." For the task factor, we explored what brain regions revealed higher functional activation during explicit compared with implicit processing and vice versa. For the factor emotion, we were interested in increased brain activations elicited by angry compared with neutral stimuli. We first compared the general brain activity of angry with neutral stimuli across both tasks. Moreover, the same contrast was also computed separately for the explicit and implicit task to find out how the general brain network for the processing of angry compared with neutral prosody emerges during both tasks. Finally, to find out brain activity that is exclusive to the processing of angry prosody during a specific task, we finally also did 2 different interaction analyses including the combination of angry prosody with the explicit and the implicit task, respectively. All contrasts were thresholded at $P < 0.001$ (uncorrected). Cluster extend threshold corrected for multiple comparison was computed based on the estimated smoothness of the data according to Forman et al. (1995) as implemented in the Brain Voyager QX software (version 2.2.1.1650; Brain Innovation, Maastricht, The Netherlands). An iterative Monte Carlo simulation with 1000 repetitions for each single contrast yielded a minimum cluster extent of $k = 6$ voxels corresponding to a cluster-level false-positive rate of $P < 0.001$.

We were additionally interested in the influence of basic auditory stimulus features on functional brain activations. We therefore included the log-transformed mean and variation (SD) of the F0 and the energy (I) of each auditory stimulus in 2 additional analyses. In a first analysis, we included the mean and SD of the F0 as a covariate on a trial-by-trial basis in the first-level analysis. This analysis should reveal functional activations that are insensitive to differences in F0 stimulus features. The same analysis was repeated by including the mean and SD of I as a covariate. All contrasts were thresholded at $P < 0.001$ (uncorrected) with a cluster extent of $k = 6$ voxels.

For the voice localizer, we finally contrasted vocal against nonvocal animal and environmental stimuli at a threshold of $P < 0.001$

(uncorrected) and a cluster extent of $k = 6$ voxels. We determined voice-sensitive regions along the STG and STS in both hemispheres for each participant as well as for the entire sample.

Results

Regions in the superior temporal cortex are not only sensitive to emotional prosody but more generally to human voices (Belin et al. 2000). To define voice-sensitive areas in superior temporal cortex, we first ran a voice localizer fMRI scan where participants listened to either 8-s auditory stimuli of human nonspeech voices or to auditory stimuli of environmental sounds or animal vocalizations (see http://vnl.psy.gla.ac.uk/ and Belin et al. 2000). We compared functional activations for human voices with activations for environmental and animal sounds and revealed extended bilateral activations in the STG and STS for the entire group of participants (see Supplementary Fig. 1 and Supplementary Table 1). These group activations served to define voice-sensitive areas in the bilateral temporal cortex (black outline in fig2, 3, 5, and 6) and were confirmed by single-subject analysis on the voice-sensitive regions (see Supplementary Fig. 1B). Our results of voice-sensitive areas are consistent with a recent study, which has shown a differential decoding of emotional prosody within these voice-sensitive regions (Ethofer, Van De Ville, et al. 2009).

Frontotemporal and Subcortical Subregions for the Decoding of Emotional Prosody
To reveal voice-sensitive subregions in superior temporal cortex as well as subregions in inferior frontal and subcortical brain regions that are involved in the decoding of emotional prosody, we asked 17 participants to perform a prosody (explicit) or a gender discrimination task (implicit) on angry and neutral voices, while we collected another set of functional fMRI scans (Fig. 1A). Functional scans were aligned to the posterior–anterior orientation of STS and were spatially restricted to superior temporal cortex and inferior frontal cortex covering also subcortical regions such as the amygdala to obtain high-resolution images from these brain regions. We used a sparse temporal acquisitions protocol where auditory
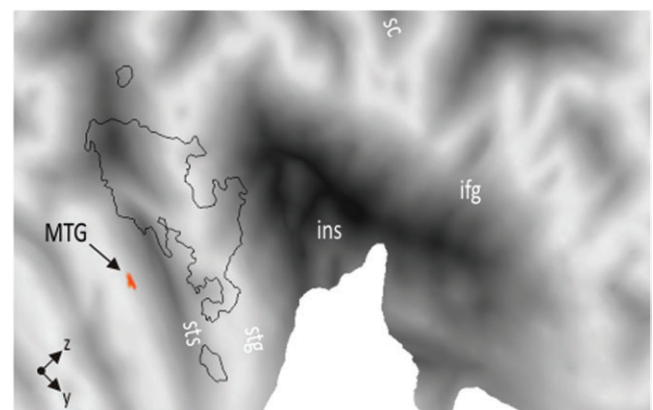

Figure 2. Comparing functional activations during the gender discrimination task with the prosody discrimination task irrespective of the emotional value of the voice revealed activity in right MTG. The figure shows functional modulations of activity rendered on a flat representation of the right hemispheric cortical surface of the human Colin atlas implemented in CARET software (Van Essen et al. 2001). ifg, inferior frontal gyrus; ins, insula; MTG, middle temporal gyrus; stg, superior temporal gyrus; sts, superior temporal sulcus.
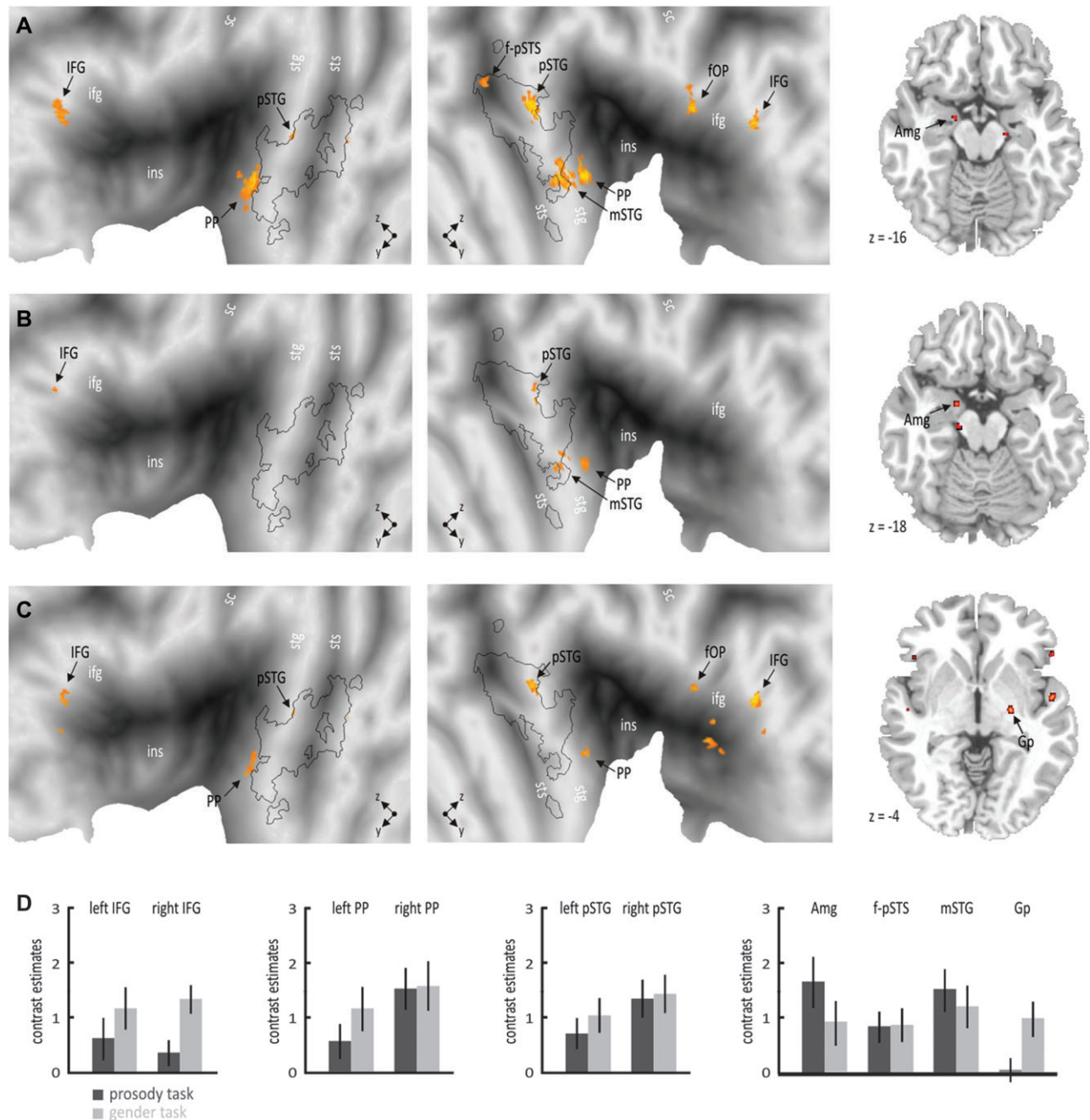
Figure 3. Increased functional activations for angry compared with neutral voices (A) independent of the task (upper row), (B) during the explicit prosody discrimination task (middle row), and (C) during the implicit gender discrimination task (lower row). The black outline represents the voice-sensitive areas as revealed by the voice localizer scan (see Supplementary Fig. 1). Many regions in the superior temporal and inferior frontal cortex were generally active independent of the task (right f-pSTS) or during both the explicit and the implicit task (left IFG). The amygdala was active only during the explicit task, whereas other regions became active only during the implicit task, such as the right IFG and globus pallidus, as well as the left PP and pSTG. (D) Contrast estimates for selected regions for the comparison of angry with neutral prosody separately for the prosody discrimination task (dark gray bars) and the gender discrimination task (light gray bars); error bars denote the SEM. Lowercase abbreviations indicate anatomical regions and uppercase abbreviations indicate functional activations. Amg, amygdala; Cd, caudate nucleus; fOP, frontal operculum; f-pSTS, fundus of the posterior superior temporal sulcus; Gp, globus pallidus; HG, Heschl gyrus; IFG/ifg, inferior frontal gyrus; ins, insula; mSTG, midsuperior temporal gyrus; PP, planum polare; pSTG, posterior superior temporal gyrus; PT, temporal plane; Put, putamen; sc, central sulcus; sgACC, subgenual anterior cingulate gyrus; stg, superior temporal gyrus; sts, superior temporal sulcus.

stimuli were presented in an 8.25-s silent gap between image acquisitions (see Materials and Methods).

In a first analysis, we explored the effects of task by comparing functional activations during the prosody discrimination task (explicit processing) with activations during the gender discrimination task (implicit processing) irrespective of the emotional value of the voice. While explicit compared with implicit processing of prosody revealed no functional activation in our primary regions of interest in lateral superior temporal and inferior frontal lobe (see Supplementary Table 2A for a full

list of activations), implicit compared with explicit processing elicited activity in right middle temporal gyrus (MTG, MNI x, y, z; 66, −7, −16; Fig. 2).

Besides the main effects of task, we were furthermore interested in the general effects of angry compared with neutral voices across both tasks. This analysis should reveal brain activity, which is independent of the level of processing. We found that several subregions in superior temporal and inferior cortex are sensitive to emotional prosody. Especially, in the right hemisphere, we found temporal subregions in the fundus of the posterior superior temporal gyrus (f-pSTS; 45, −34, 4), pSTG (69, −22, 4), mSTG (66, −3, 2), and planum polare (PP; 53, −4, −4) as well as frontal subregions in the frontal operculum (fOP; 48, 13, −2) and IFG (51, 32, −2). The fOP activation was located in the right hemispheric homologue to left BA 44, whereas the other IFG activation was located on BA 45/47. In the left hemisphere, we found activity in PP (−50, −10, 4), pSTG (−68, −27, 6), IFG (−44, 29, 0) as well as in the amygdala (−18, −4, −16). All subregions in the superior temporal cortex were located in voice-sensitive areas, except for the activity in PP (Fig. 3A; see Supplementary Table 2C for a full list of activations).

Since, however, activity in this frontotemporal and subcortical set of brain regions might also depend on the level of processing, we additionally compared activity for angry compared with neutral voices separately for the explicit and the implicit discrimination task. We found the left IFG (−50, 26, 6) and amygdala (−18, −4, −18), as well as the right PP (53, −4, −4), mSTG (60, 1, 0), and pSTG (65, −24, 6), to be active only during the explicit discrimination task (Fig. 3B; see Supplementary Table 2D), whereas the left IFG (−42, 29, 0) and PP (−51, −10, −4), as well as the right IFG (54, 31, 2), fOP (47, 13, 0), PP (53, −4, −4), pSTG (68, −22, 4), and globus pallidus (23, −13, 4), were active during the implicit discrimination task (Fig. 3C; see Supplementary Table 2E). These results suggest that some regions in STG and IFG are generally sensitive to emotional prosody independent of the level of processing (right f-pSTG, left IFG), while other regions are active only during the explicit (amygdala) or implicit decoding of emotional prosody (right IFG, left PP, left pSTG). To further specify specific brain activity during the explicit or implicit discrimination task, we performed an interaction analysis to find brain activity, which is unique for the decoding of angry prosody during a specific level of processing. This analysis revealed specific activations for angry voices in the explicit discrimination task in the subgenual anterior cingulate cortex (sgACC; −5, 31, −10) and in the left striatum putamen (Fig. 4A) and bilateral striatal activity during the implicit discrimination task (Fig. 4B; see Supplementary Tables 2F,G).

Taken together, we found a general frontotemporal and subcortical network of brain regions for the decoding of angry prosody that consist of several local brain regions especially in the right hemisphere. Activity in this general set of brain regions, however, revealed a dependency on the levels of processing. Another factor, which is assumed to influence activity in this frontotemporal network for the sensory and evaluative decoding of emotional prosody are emotion-specific basic acoustical features, such as the F0 and the energy of stimuli, which are known to vary across different vocal expressions of emotions (Banse and Scherer 1996; Grandjean et al. 2006).

Sensitivity to Emotion-Specific Acoustic Features in posterior STG and Amygdala

We performed 2 additional analyses that were similar to the one described in the former section but taking into account F0 and energy (I) differences between angry and neutral stimuli. Specifically, we scored the mean and SD of the F0 and the energy for each of the 16 angry and 16 neutral auditory stimuli. These stimulus features were log transformed and entered into the statistical analysis for each participant on a trial-by-trial basis (angry voices: $\log F0_{mean}$ = 5.83 [standard error of the mean—SEM 0.04], $\log F0_{SD}$ = 3.74 [SEM 0.15], $\log I_{mean}$ = 4.29 [SEM 0.01], $\log I_{SD}$ = 1.93 [SEM 0.07]; neutral voices: $\log F0_{mean}$ = 5.00 [SEM 0.10], $\log F0_{SD}$ = 2.66 [SEM 0.20], $\log I_{mean}$ = 4.31 [SEM 0.01], $\log I_{SD}$ = 1.66 [SEM 0.06]). We performed one analysis by taking into account the mean and SD of the F0 and a separate analysis with the mean and SD of the energy. Group analyses and contrasts were performed in the same way as described in the former section. Functional activations, which are modulated by this analysis compared with our former analysis, should indicate a strong sensitivity to acoustical stimulus features.

When including the mean and SD of the F0 as covariate in the analysis, the right MTG activations that we found during the gender compared with the prosody discrimination task remained active indicating insensitive to F0 stimulus features (see Supplementary Table 3B). We also found more activity in the right f-pSTS (47, −33, 0) as well as in left aSTG (−56, 11, −10) and IFG (−45, 41, −10) for the prosody compared with gender discrimination task (Fig. 5; see Supplementary Table 3A). For the contrasts related the experimental factor emotion, most of the activations for angry compared with neutral voices in subcortical regions and in posterior regions of STG disappeared, except for the activity in right mSTG (59, −1, −2) and PP (51, −4, 0), which seems also insensitive to F0 stimulus features (Fig. 6; see Supplementary Table 3C–E). Furthermore, activity in
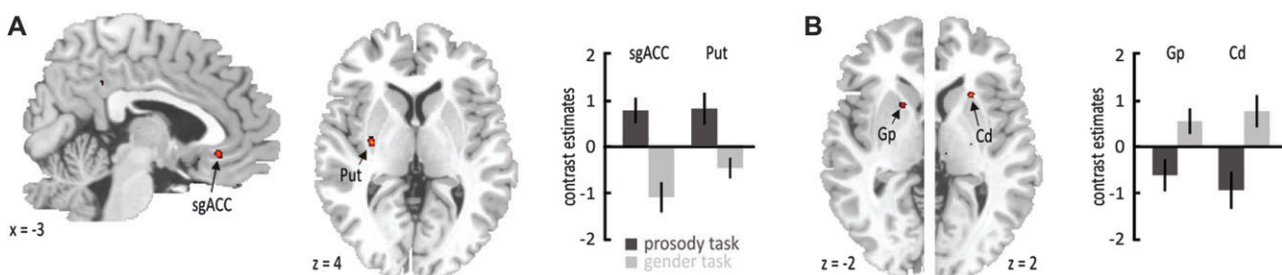

Figure 4. Specific functional activations as revealed by the interaction analysis for angry compared with neutral voices during (A) the explicit prosody categorizations task in the sgACC and (B) during the implicit gender decision task in the left globus pallidus (Gp) and right caudate nucleus (Cd).
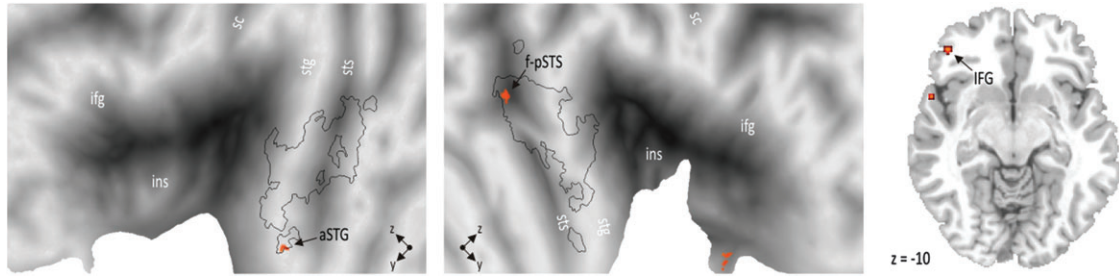
Figure 5. Prosody discrimination compared with gender discrimination revealed functional activity in right f-pSTS as well as left aSTG, and IFG when taking mean and SD of the F0 into account. Similar results were found when taking mean and SD of the energy into account (see Supplementary Fig. 2). For abbreviations, see Figure 3.
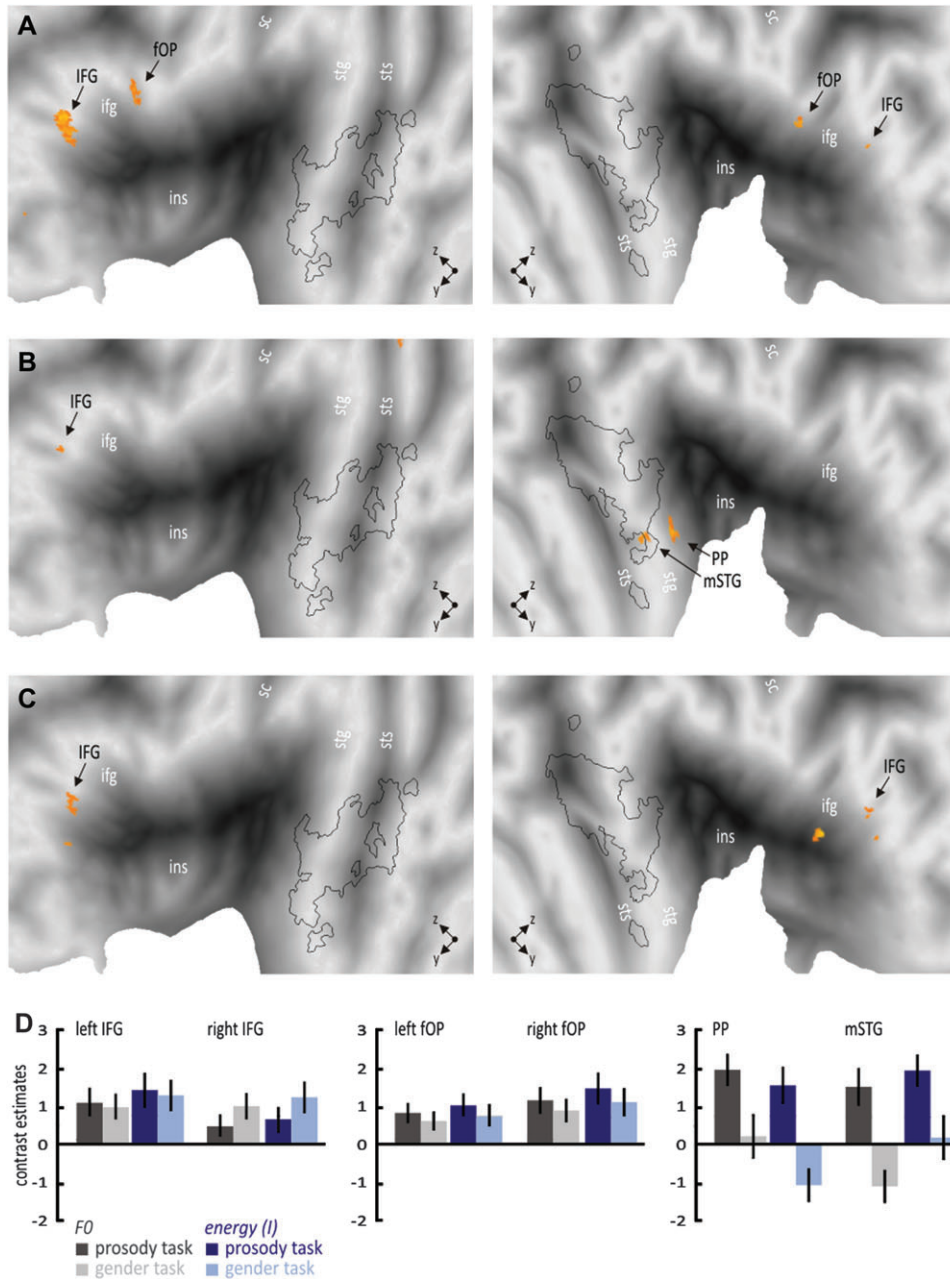

Figure 6. Functional activations for angry compared with neutral voices, taking into account the mean and the SD of the F0. Functional activations are shown (*A*) independent of the task (upper row); (*B*) during the explicit prosody categorizations task (middle row); and (*C*) during the implicit gender decision task (lower row). We obtained an almost identical pattern of activations when taking the mean and SD of the energy (*I*) into account (see Supplementary Fig. 2). (*D*) Contrast estimates for activated regions for the analysis with the F0 (gray bars) and with *I* (blue bars) separately for the prosody task (dark bars) and for the gender task (light bars). For abbreviations, see Figure 3.

bilateral IFG also showed an insensitivity to F0 stimulus features, where left IFG (–51, 26, 4) again was generally active during both tasks, whereas right IFG (51, 32, –2) was only active during implicit processing. We again explored specific activations using an interaction analysis to find unique activation for angry prosody during explicit or implicit processing. Compared with our first analysis that did not take F0 stimulus features into account, we found right globus pallidus (23, –15, 0) activity during implicit processing (Fig. 7D; see Supplementary Table 3G) and widespread activations in bilateral mSTG (left; –62, –1, –4 and right; 60, 1, 0) and pSTG (left; –65, –33, 4 and right; 65, –22, 6) as well as in the right fOP (56, 8, 14), Heschl gyrus (HG; 65, –16, 12), and planum temporale (PT; 50, –22, 4) for angry voices during the explicit task (Fig. 7A–C; see Supplementary Table 3F). Especially during explicit processing, additionally, activity was found in region in between right mSTG and pSTG (termed as m-pSTG; –60, –18, 2) located within voice-sensitive superior temporal cortex.

We performed an identical analysis with the mean and SD of the energy as covariate and found almost identical activations compared with the analysis with the mean and SD of the F0 as covariate (see Supplementary Table 4 and Supplementary Figs 2–4). The only exception was additional activation in the sgACC (–5, 31, –10) for the interaction analysis including angry voices in the explicit task (see Supplementary Table 4F), suggesting that sgACC is less sensitive to energy differences than to F0 differences. However, the general similarity for the analysis with F0 and the energy as covariate indicates a comparable influence of F0 and energy stimulus features on frontotemporal and subcortical brain activity since F0 and energy features seem to be highly associated during angry prosody (Banse and Scherer 1996; Leitman et al. 2010). This influence of F0 and energy stimulus features was especially pronounced in pSTG and m-pSTG regions and especially when attention was explicitly focused on the emotional prosody and most probably on prosody-specific acoustical stimulus features.

Discussion

The results of present study suggest that the frontal, temporal, and subcortical network that is commonly involved in the decoding of emotional prosody consists of several subregions especially at the cortical level. We found that this differentiation in superior temporal and inferior frontal brain regions was strongly driven by the comparison between angry and neutral prosody, while the difference between the prosody and the gender discrimination task did not reveal strong differences except for an activity in right MTG during the gender
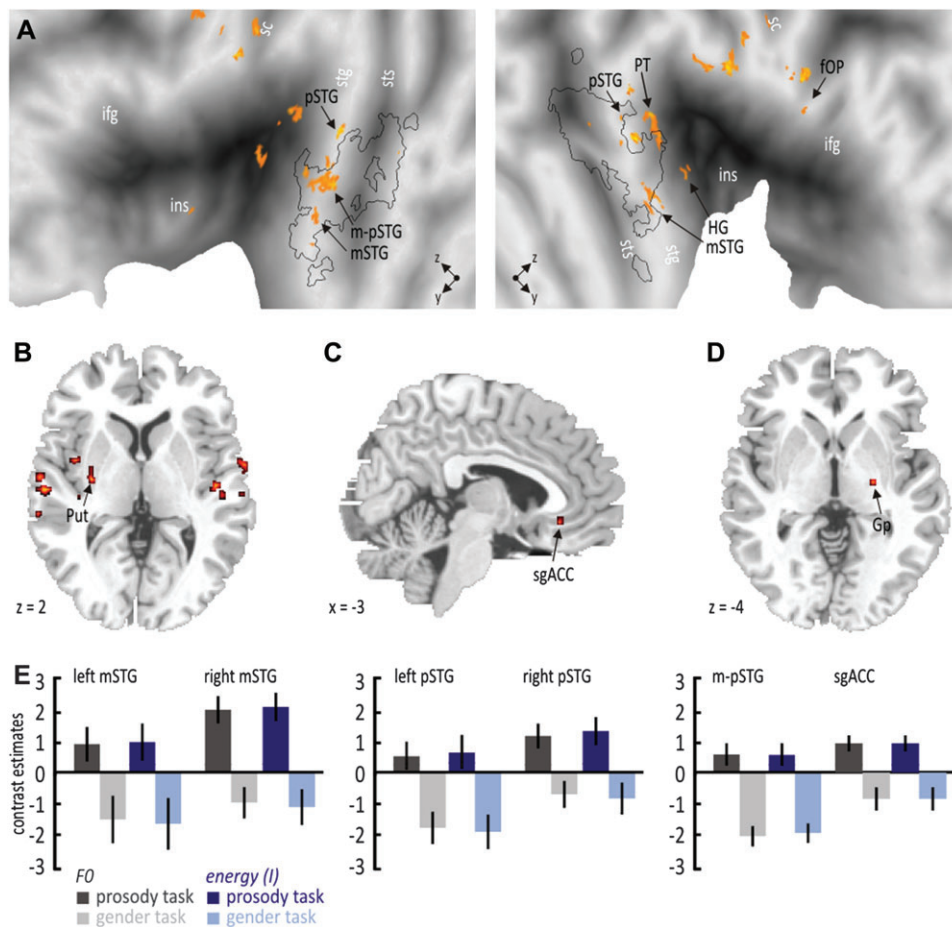

Figure 7. Specific activations as revealed by the interaction analysis for angry compared with neutral voices, taking into account the mean and the SD of the F0 and of the energy (*I*). Both analyses with the F0 and the *I* as covariate revealed an almost identical pattern of results. (*A*) Angry voices during the prosody task revealed bilateral functional activations along the STG and in the right regions of the PT and the HG of the STG as well as (*B*) in the left putamen. (*C*) A specific activation for angry voices during the prosody task was found in the sgACC but only with *I* as a covariate. (*D*) Angry voices during the gender task revealed specific activations in the right globus pallidus, both with F0 and *I* as a covariate. (*E*) Contrast estimates for selected regions. For abbreviations, see Figure 3.

discrimination task. This might be indicative of increased gender encoding as part of a general speaker's identity processing in more anterior temporal brain regions (Belin and Zatorre 2003; Formisano et al. 2008). The fact that we did not find strong effects for the explicit prosody discrimination task might be due to our restricted scanning space mainly including superior temporal and inferior frontal brain regions, while recent studies mainly found activity in superior parietal and superior frontal activity during the explicit decoding of emotional prosody (Bach et al. 2008; Ethofer, Kreifelts, et al. 2009). However, only when taking central prosodic stimulus features into account, we revealed a set of distributed brain regions during explicit decoding of emotional prosody, as we will discuss below.

Though task factor alone was not able to reveal a strong differentiation on brain regions, it showed a strong influence when combined with the factor emotion. The comparison of angry with neutral prosody revealed several temporal and frontal subregions, which become differentially active during task-independent, explicit, or implicit decoding of emotional compared with neutral prosody. The right hemisphere in particular revealed at least a temporal network of 4 regions, namely, regions in the fundus of the superior temporal sulcus (f-pSTS), pSTG, mSTG, and PP. All these regions were located within voice-sensitive areas, except for the PP. The f-pSTS has not frequently been reported as a region in the superior frontal cortex that is sensitive to emotional prosody but might correspond with posterior STS regions involved in the integration of multimodal emotional signals (Kreifelts et al. 2009). The f-pSTS was active independent of the task, whereas the pSTG and PP were similarly active during both tasks, and the mSTG was active only during explicit prosody discrimination. This posterior-to-anterior gradient for a task-independent and stimulus-driven decoding of emotional prosody to more explicit decoding in the superior temporal cortex is in accordance with a former observation (Grandjean et al. 2005) and resembles a proposed increase in levels of stimulus processing when information is fed forward to the more anterior superior temporal cortex (Schirmer and Kotz 2006) along a proposed pathway of auditory object recognition (Rauschecker and Scott 2009).

This temporal network showed a strong differentiation in the right hemisphere, whereas only 2 regions could be differentiated in the left hemisphere. A region in the left pSTG located in voice-sensitive areas was active during implicit processing and PP during both implicit and explicit processing. Although this left hemispheric pattern of activations follows the same posterior-to-anterior gradient, the right hemispheric dominance is in accordance with a stronger, but not exclusive, involvement of the right superior temporal cortex in decoding emotional prosody of vocalizations consisting of nonintelligible speech (Grandjean et al. 2005; Sander et al. 2005) or nonspeech stimuli (Fecteau et al. 2007).

Similar to the differentiation of functional activations in the superior temporal cortex, peak activations in the inferior frontal cortex could be located in different subregions and also showed task dependency. Angry compared with neutral voices elicited increased activity in the right fOP and IFG during implicit processing and in the left IFG during both implicit and explicit processing. The distinction of activation in the right inferior frontal cortex supports the notion that different subregions might subserve different functional roles. Activity

in the fOP (BA 45) might subserve increased cognitive evaluation of emotional prosody (Leitman et al. 2010), whereas more anterior IFG regions (BA 47) are associated with outcome-related evaluations (Schirmer and Kotz 2006). Both processes seem to be especially increased during implicit decoding when attentional focus is not directly focused on the emotional prosody feature. Explicit decoding of emotional prosody engaged only the left IFG in BA47 and is in accordance with recent studies (Bach et al. 2008; Ethofer, Kreifelts, et al. 2009) and with studies showing increased left IFG activations when explicit decoding is stressed by contextual factors (Schirmer et al. 2004; Mitchell 2006). Specific activity during explicit decoding of emotional prosody was also found in the ventromedial frontal cortex and might serve in similar elaborate decoding and appraisal processes of emotional prosody when individuals attend to emotional prosodic stimulus features (Sander et al. 2005).

Explicit decoding of emotional prosody also revealed activation in the left amygdala. Some recent studies also report activation in the amygdala for emotional compared with neutral voices (Sander et al. 2005; Fecteau et al. 2007; Bach et al. 2008; Ethofer, Kreifelts, et al. 2009; Wiethoff et al. 2009). However, though some studies report amygdala activations only during explicit processing of emotional prosody (Wiethoff et al. 2009; Leitman et al. 2010), other studies highlight the notion that the amygdala is more generally active independent of the level of processing (Sander et al. 2005) or specifically during the implicit processing of emotional prosody (Bach et al. 2008). In the latter case, the amygdala is supposed to act as a detector of important emotional information even when this information is presented outside the focus of attention. We also found, in accordance with the results of Sander et al. (2005), general left amygdala activation independent of the task. However, Bach et al. (2008) found amygdala activation during implicit processing only when comparing all emotional and neutral stimuli together with the explicit processing of both kinds of stimuli. In the present study, we specifically compared emotional and neutral stimuli within each task separately and did not find specific amygdala activity for emotional stimuli during implicit processing; rather, we found this activity during explicit processing.

Apart from activity in the amygdala, which might code the emotional value of emotional prosody during explicit processing, we found specific activity in the left basal ganglia during explicit decoding but also in the bilateral basal ganglia during implicit decoding of emotional prosody. Rather than coding the emotional value, the basal ganglia are assumed to code the temporal patterns of emotional acoustic cues, such as rhythms or variations in the auditory signal (Kotz and Schwartze 2010). Angry as compared with neutral prosody is especially characterized by a strong variation of the F0 and the energy, rapid speech onset and high speech rate that results in a specific and distinguished temporal acoustical pattern (Banse and Scherer 1996; Grandjean et al. 2006).

Whereas basal ganglia seem to specifically code the temporal pattern of emotional acoustic cues, cortical brain regions seem to be sensitive to other acoustic cues (Wiethoff et al. 2008; Leitman et al. 2010). We tested whether the different temporal and frontal subregions that we found to be active during the explicit and implicit decoding of emotional prosody are sensitive to F0 and energy stimulus features of the angry compared with neutral prosody. When including mean and

variation of the F0 or of the energy as a covariate in our analysis, we found that bilateral activations in the inferior cortex were insensitive to these acoustic stimulus features. Moreover, during the prosody discrimination task, we found additional activity in a left anterior IFG region (Ethofer, Anders, Erb, et al. 2006; Bach et al. 2008), which might represent a frontal voice and prosody sensitive area (Fecteau et al. 2005). This left anterior IFG regions seem not only insensitive to acoustic features of emotional prosody but also seem to be covered by F0 and energy differences between conditions in the first analysis. Explicit attention to emotional prosody also revealed additional activity in a right posterior and left anterior voice-sensitive area, indicating that the explicit decoding of emotional prosody in general includes a temporofrontal network that can only be detected when central prosody features are included in the analysis, for which a different set of brain regions seems to be sensitive.

For the comparison of emotional with neutral prosody, we found that especially posterior regions in STG are sensitive to prosodic stimulus features. Except for activations in the right PP and mSTG, activations in the amygdala, the sgACC (only for the F0 as covariate), and most of the activations in the bilateral STG disappeared, indicating sensitivity to acoustic stimulus features. However, the right PP and mSTG were active only during explicit processing, which we found even when taking acoustical features into account. These activations might indicate some higher level auditory emotional representations independent of more basic stimulus features but depending on explicit attention to the emotional prosody. Furthermore, the interaction analysis of specific activations revealed widespread bilateral STG activations, again, especially during explicit decoding of angry prosody. We found activations in a left region in between the mSTG and pSTG (m-pSTG) as well as in the right HG and PT. Explicitly orienting attention to the prosodic features and presumably to emotion-specific acoustical features of angry compared with neutral stimuli might have led to this enhanced activity in these regions. These regions were not active during the first analysis, in which we did not take stimulus features into account. Therefore, these activations again might have been covered by strong F0 and energy differences between conditions. In particular, the m-pSTG was located in more pSTG regions for implicit decoding and in mSTG regions for explicit decoding. This finding complements the formerly discussed posterior-to-anterior gradient of voice-sensitive STG regions, where a more anterior location implies a greater sensitivity to emotional prosody during explicit decoding.

We obtained similar results when including mean and variation of the energy as a covariate in the analysis instead of the F0, the only exception being that the sgACC showed no sensitivity to energy-related stimulus features. The similarity of results could be based on the fact that F0 and energy modulations are highly correlated for angry expressions and can separate anger from other emotions (Banse and Scherer 1996; Leitman et al. 2010; Patel et al. 2011). Though it is assumed that the left hemisphere is usually more sensitive to the F0, whereas the right hemisphere is sensitive to energy (Zatorre et al. 2002), we could not confirm this hemispheric specialization. Regions in the bilateral STG and amygdala showed a similar sensitivity to F0 in addition to energy-related stimulus features.

A surprising finding was the sensitivity of amygdala activation to F0 and energy stimulus features. Activity in the amygdala is usually assumed to reflect the decoding of the emotional value of auditory stimuli rather than the decoding of basic acoustic cues (Scott et al. 1997; Sprengelmeyer et al. 1999; Wiethoff et al. 2009; Leitman et al. 2010). However, Bordi and LeDoux (1992) have shown a sensitivity of the amygdala to simple auditory sensory stimulation suggesting some sensitivity of the amygdala to basic acoustic stimulus features. Angry prosody consists of a unique combination of F0 and energy stimulus features, and activity in the amygdala in response to angry prosody (Grandjean et al. 2005; Sander et al. 2005) might partly rely on these basic acoustic stimulus cues (Leitman et al. 2010).

Taken together, the present data revealed 3 major findings. First, the common frontotemporal network of brain regions consists of several subregions. We were able to distinguish at least 4 subregions in the right STG and 2 subregions in the IFG when comparing emotional to neutral prosody but also during the prosody compared with gender discrimination task when taking stimulus features into account. Second, these subregions are differentially sensitive to the levels of processing. Implicit processing of emotional compare with neutral prosody engages more pSTG regions, the bilateral IFG, and bilateral basal ganglia, whereas explicit processing relies on more mSTG regions, the left IFG, amygdala, left basal ganglia, and sgACC. Third, a part of these regions also showed sensitivity to emotion-specific acoustic cues when comparing emotional with neutral prosody. This sensitivity was specifically strong in the bilateral pSTG during implicit processing, whereas explicit processing revealed a widespread network of bilateral mSTG and pSTG regions, which were relatively independent of F0 and energy-related acoustic cues.

Supplementary Material

Supplementary material can be found at: http://www.cercor. oxfordjournals.org/

Funding

Notes

References

Adolphs R, Tranel D. 1999. Intact recognition of emotional prosody following amygdala damage. Neuropsychologia. 37:1285–1292.

Anderson AK, Phelps EA. 2002. Is the human amygdala critical for the subjective experience of emotion? Evidence of intact dispositional affect in patients with amygdala lesions. J Cogn Neurosci. 14:709–720.

Bach DR, Grandjean D, Sander D, Herdener M, Strik WK, Seifritz E. 2008. The effect of appraisal level on processing of emotional prosody in meaningless speech. Neuroimage. 42:919–927.

Banse R, Scherer KR. 1996. Acoustic profiles in vocal emotion expression. J Pers Soc Psychol. 70:614–636.

Bänziger T, Scherer KR. 2010. Introducing the Geneva Multimodal Emotion Portrayal (GEMEP) corpus. In: Bänziger T, Scherer KR, Roesch EB, editors. Blueprint for affective computing: a sourcebook. Oxford: Oxford University Press. p. 271–294.

Beaucousin V, Lacheret A, Turbelin MR, Morel M, Mazoyer B, Tzourio-Mazoyer N. 2007. FMRI study of emotional speech comprehension. Cereb Cortex. 17:339–352.

Belin P, Fecteau S, Charest I, Nicastro N, Hauser MD, Armony JL. 2008. Human cerebral response to animal affective vocalizations. Proc Biol Sci. 275:473–481.

Belin P, Zatorre RJ. 2003. Adaptation to speaker's voice in right anterior temporal lobe. Neuroreport. 14:2105–2109.

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

Bordi F, LeDoux J. 1992. Sensory tuning beyond the sensory system: an initial analysis of auditory response properties of neurons in the lateral amygdaloid nucleus and overlying areas of the striatum. J Neurosci. 12:2493–2503.

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

Calder AJ, Keane J, Lawrence AD, Manes F. 2004. Impaired recognition of anger following damage to the ventral striatum. Brain. 127:1958–1969.

Ethofer T, Anders S, Erb M, Herbert C, Wiethoff S, Kissler J, Grodd W, Wildgruber D. 2006. Cerebral pathways in processing of affective prosody: a dynamic causal modeling study. Neuroimage. 30:580–587.

Ethofer T, Anders S, Wiethoff S, Erb M, Herbert C, Saur R, Grodd W, Wildgruber D. 2006. Effects of prosodic emotional intensity on activation of associative auditory cortex. Neuroreport. 17:249–253.

Ethofer T, Kreifelts B, Wiethoff S, Wolf J, Grodd W, Vuilleumier P, Wildgruber D. 2009. Differential influences of emotion, task, and novelty on brain regions underlying the processing of speech melody. J Cogn Neurosci. 21:1255–1268.

Ethofer T, Van De Ville D, Scherer K, Vuilleumier P. 2009. Decoding of emotional information in voice-sensitive cortices. Curr Biol. 19:1028–1033.

Fecteau S, Armony JL, Joanette Y, Belin P. 2005. Sensitivity to voice in human prefrontal cortex. J Neurophysiol. 94:2251–2254.

Fecteau S, Belin P, Joanette Y, Armony JL. 2007. Amygdala responses to nonlinguistic emotional vocalizations. Neuroimage. 36:480–487.

Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC. 1995. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. Magn Reson Med. 33:636–647.

Formisano E, De Martino F, Bonte M, Goebel R. 2008. "Who" is saying "what"? Brain-based decoding of human voice and speech. Science. 322:970–973.

Grandjean D, Banziger T, Scherer KR. 2006. Intonation as an interface between language and affect. Prog Brain Res. 156:235–247.

Grandjean D, Sander D, Lucas N, Scherer KR, Vuilleumier P. 2008. Effects of emotional prosody on auditory extinction for voices in patients with spatial neglect. Neuropsychologia. 46:487–496.

Grandjean D, Sander D, Pourtois G, Schwartz S, Seghier ML, Scherer KR, Vuilleumier P. 2005. The voices of wrath: brain responses to angry prosody in meaningless speech. Nat Neurosci. 8:145–146.

Kotz SA, Meyer M, Alter K, Besson M, von Cramon DY, Friederici AD. 2003. On the lateralization of emotional prosody: an event-related functional MR investigation. Brain Lang. 86:366–376.

Kotz SA, Schwartze M. 2010. Cortical speech processing unplugged: a timely subcortico-cortical framework. Trends Cogn Sci. 14:392–399.

Kreifelts B, Ethofer T, Shiozawa T, Grodd W, Wildgruber D. 2009. Cerebral representation of non-verbal emotional perception: fMRI reveals audiovisual integration area between voice- and face-sensitive regions in the superior temporal sulcus. Neuropsychologia. 47:3059–3066.

Lee T-W, Josephs O, Dolan RJ, Critchley HD. 2006. Imitating expressions: emotion-specific neural substrates in facial mimicry. Soc Cogn Affect Neurosci. 1:122–135.

Leitman DI, Wolf DH, Ragland JD, Laukka P, Loughead J, Valdez JN, Javitt DC, Turetsky BI, Gur RC. 2010. "It's not what you say, but how you say it": a reciprocal temporo-frontal network for affective prosody. Front Hum Neurosci. 4:1–13.

Mitchell RL. 2006. How does the brain mediate interpretation of incongruent auditory emotions? The neural response to prosody in the presence of conflicting lexico-semantic cues. Eur J Neurosci. 24:3611–3618.

Mitchell RL, Boucas SB. 2009. Decoding emotional prosody in Parkinson's disease and its potential neuropsychological basis. J Clin Exp Neuropsychol. 31:553–564.

Mitchell RL, Elliott R, Barry M, Cruttenden A, Woodruff PW. 2003. The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. Neuropsychologia. 41:1410–1421.

Nakamura K, Kawashima R, Ito K, Sugiura M, Kato T, Nakamura A, Hatano K, Nagumo S, Kubota K, Fukuda H, et al. 1999. Activation of the right inferior frontal cortex during assessment of facial emotion. J Neurophysiol. 82:1610–1614.

Patel S, Scherer KR, Bjorkner E, Sundberg J. 2011. Mapping emotions into acoustic space: the role of voice production. Biol Psychol. 87:93–98.

Rauschecker JP, Scott SK. 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nat Neurosci. 12:718–724.

Sander D, Grandjean D, Pourtois G, Schwartz S, Seghier ML, Scherer KR, Vuilleumier P. 2005. Emotion and attention interactions in social cognition: brain regions involved in processing anger prosody. Neuroimage. 28:848–858.

Schirmer A, Kotz SA. 2006. Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. Trends Cogn Sci. 10:24–30.

Schirmer A, Zysset S, Kotz SA, Yves von Cramon D. 2004. Gender differences in the activation of inferior frontal cortex during emotional speech perception. Neuroimage. 21:1114–1123.

Scott SK, Young AW, Calder AJ, Hellawell DJ, Aggleton JP, Johnson M. 1997. Impaired auditory recognition of fear and anger following bilateral amygdala lesions. Nature. 385:254–257.

Sprengelmeyer R, Young AW, Schroeder U, Grossenbacher PG, Federlein J, Buttner T, Przuntek H. 1999. Knowing no fear. Proc Biol Sci. 266:2451–2456.

Van Essen DC, Drury HA, Dickson J, Harwell J, Hanlon D, Anderson CH. 2001. An integrated software suite for surface-based analyses of cerebral cortex. J Am Med Inform Assoc. 8:443–459.

Wiethoff S, Wildgruber D, Grodd W, Ethofer T. 2009. Response and habituation of the amygdala during processing of emotional prosody. Neuroreport. 20:1356–1360.

Wiethoff S, Wildgruber D, Kreifelts B, Becker H, Herbert C, Grodd W, Ethofer T. 2008. Cerebral processing of emotional prosody—influence of acoustic parameters and arousal. Neuroimage. 39:885–893.

Wildgruber D, Ethofer T, Grandjean D, Kreifelts B. 2009. A cerebral network model of speech prosody comprehension. Int J Speech Lang Pathol. 11:277–281.

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.

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, Riecker A, Hertrich I, Erb M, Grodd W, Ethofer T, Ackermann H. 2005. Identification of emotional intonation evaluated by fMRI. Neuroimage. 24:1233–1241.

Zatorre RJ, Belin P, Penhune VB. 2002. Structure and function of auditory cortex: music and speech. Trends Cogn Sci. 6:37–46.