The effect of appraisal level on processing of emotional prosody in meaningless speech

Dominik R. Bach [a,b,*], Didier Grandjean [c,d], David Sander [c,d], Marcus Herdener [a], Werner K. Strik [a], Erich Seifritz [a,e]

[a] University Hospital of Psychiatry, University of Bern, Switzerland
[b] Wellcome Trust Centre for Neuroimaging, University College London, UK
[c] Swiss Center for Affective Sciences, University of Geneva, Switzerland
[d] Department of Psychology, University of Geneva, Switzerland
[e] Psychiatric Private Clinic Sanatorium Kilchberg, Kilchberg, Switzerland

ARTICLE INFO

ABSTRACT

In visual perception of emotional stimuli, low- and high-level appraisal processes have been found to engage different neural structures. Beyond emotional facial expression, emotional prosody is an important auditory cue for social interaction. Neuroimaging studies have proposed a network for emotional prosody processing that involves a right temporal input region and explicit evaluation in bilateral prefrontal areas. However, the comparison of different appraisal levels has so far relied upon using linguistic instructions during low-level processing, which might confound effects of processing level and linguistic task. In order to circumvent this problem, we examined processing of emotional prosody in meaningless speech during gender labelling (implicit, low-level appraisal) and emotion labelling (explicit, high-level appraisal). While bilateral amygdala, left superior temporal sulcus and right parietal areas showed stronger blood oxygen level-dependent (BOLD) responses during implicit processing, areas with stronger BOLD responses during explicit processing included the left inferior frontal gyrus, bilateral parietal, anterior cingulate and supplemental motor cortex. Emotional versus neutral prosody evoked BOLD responses in right superior temporal gyrus, bilateral anterior cingulate, left inferior frontal gyrus, insula and bilateral putamen. Basal ganglia and right anterior cingulate responses to emotional versus neutral prosody were particularly pronounced during explicit processing. These results are in line with an amygdala-prefrontal-cingulate network controlling different appraisal levels, and suggest a specific role of the left inferior frontal gyrus in explicit evaluation of emotional prosody. In addition to brain areas commonly related to prosody processing, our results suggest specific functions of anterior cingulate and basal ganglia in detecting emotional prosody, particularly when explicit identification is necessary.

© 2008 Elsevier Inc. All rights reserved.

Introduction

In social interaction, another person's feelings are usually evaluated in an incidental manner. Sometimes, for example when quickly scanning a crowd of people, this might even be accomplished without conscious awareness of the individual stimuli that are being appraised. The results of this appraisal often come to one's mind only when they bear special relevance (Sander et al., 2003). On the other hand, attention can be directed in a way that allows explicit appraisal of another's feelings right away. Appraisal theories have there-fore distinguished different levels of appraisal processing: low-level appraisal, which is thought to be automatic and pre-attentive, schematic level and high-level appraisal, which is controlled and requires more cognitive resources (Leventhal and Scherer, 1987).

A number of neuroimaging studies have investigated neural networks involved in visual processing of facial affective cues. It was consistently shown that the relevant neural structures depend strongly on task instruction, and thus on the required level of appraisal processing. While facial emotion cues activate the amygdala, even when they are not consciously perceived (Killgore and Yurgelun-Todd, 2004; Whalen et al., 1998, 2004), this activity is less pronounced the more *explicit* the experimental instructions become. It should be noted here that *explicit* and *implicit* are terms in the context of a particular experimental design and can refer to different task instructions across

different studies. Thus, on different levels of cognitive demand, it was shown that amygdala activity is stronger for perceptual matching (*implicit*) than facial expression labelling (*explicit*) (Hariri et al., 2000), for gender labelling (*implicit*) than for facial expression labelling (*explicit*) (Critchley et al., 2000), and using non-facial stimuli, stronger for perceptual matching (*implicit*) than for labelling of picture type (naturalistic vs. artificial, *explicit*) (Hariri et al., 2003). On the other hand, anterior cingulate and prefrontal responses in these studies were stronger in more *explicit* compared to the *implicit* tasks, which has led to the formulation of a prefrontal-amygdala network where the prefrontal cortex exerts a top-down regulatory function over the amygdala (Hariri et al., 2003).

Beyond facial affect in the visual domain, affective prosody in the auditory domain is an important cue for inferring what another person feels (Grandjean et al., 2006). While different appraisal levels have been directly compared for processing facial affect, our knowledge regarding affective prosody is more limited. Neuroimaging studies of affective prosody investigated appraisal levels independently by examining (*a*) labelling of emotion-irrelevant cues, e.g. gender, in emotional versus neutral prosody (Grandjean et al., 2005; Sander et al., 2005), and (*b*) explicit labelling of emotional versus neutral prosody (Kotz et al., 2003), or explicit labelling of emotional prosody versus explicit labelling of emotional semantic content (Ethofer et al., 2006a,b). While different results between these types of studies might in part be related to different task instruction and appraisal level, they also used different stimulus sets, ranging from filtered speech without semantic content (Kotz et al., 2003), nonsense syllables (Grandjean et al., 2005; Sander et al., 2005), to verbal stimuli with semantic content (Ethofer et al., 2006a,b; Kotz et al., 2003).

By varying task instruction, a number of studies also implied different appraisal levels (Wambacq et al., 2004) and (*c*) directly compared explicit labelling of emotional prosody with semantic or phonetic labelling of the same stimulus (Buchanan et al., 2000; George et al., 1996; Wildgruber et al., 2004, 2005). While those studies found converging evidence for brain areas involved in explicit processing of emotional prosody versus linguistic tasks, this task design might be problematic insofar as speech-related networks are likely to be activated differently by linguistic and emotional prosody instructions. Areas with a stronger blood oxygen level-dependent (BOLD) signal during linguistic evaluation might therefore reflect implicit prosody processing as well as explicit semantic/ phonetic processing (and, on the other hand, stronger signal during explicit prosody processing might just as well be caused by implicit linguistic processing). Hence, this study design might limit the ability to distinguish the effect of emotion appraisal level and semantic/phonetic instruction. While this critique might also apply to other possible task designs (in the above cited studies on visual perception, a main effect of *implicit* emotional cue processing cannot be disentangled from, e.g. effects of gender decision), converging evidence from studies using different task instructions might eventually reveal common networks related to appraisal level rather than to specific task instructions.

Other studies on affective prosody investigated higher cognitive functions in prosody perception and included (*d*) congruent versus incongruent semantic and prosodic information (Mitchell et al., 2003), dichotic presentation and manipulation of endogenous attention (Grandjean et al., 2005; Sander et al., 2005), and audiovisual detection of emotional cues (Dolan et al., 2001; Ethofer et al., 2006c; Johnstone et al., 2006; Pourtois et al., 2005). Given the designs used, these studies can hardly contribute in assessing the effect of different appraisal levels (for a review, see Schirmer and Kotz, 2006).

Therefore, in the present study we investigated neural correlates of *explicit* and *implicit* processing of emotional cues in meaningless speech, using a non-linguistic implicit task. In addition, we aimed at assessing task × emotion interactions by including neutral stimuli into the design. Furthermore, we were interested in comparing appraisal level effects across two different emotional categories, that is, anger and fear.

Although previous studies on processing of affective prosody revealed partially inconsistent results, activity in the inferior frontal gyrus and temporal regions emerged as a common denominator. A hierarchical model has been proposed to explain these findings, where emotionally significant acoustic parameters are processed in the superior temporal gyrus (STG) and sulcus (STS), and evaluative judgements take part in the right inferior frontal and orbitofrontal gyri, while semantic processing is carried out in the left inferior frontal gyrus (Schirmer and Kotz, 2006). A recent dynamic causal modelling study has found evidence for a model where right temporal areas serve as input region for affective prosody, while further processing is accomplished in the bilateral inferior frontal gyrus (Ethofer et al., 2006a).

Thus, we aimed at testing a part of this model by directly comparing *explicit* and *implicit* processing of emotional prosody. We hypothesized that basic auditory features of the stimuli would be processed under both task instructions, while evaluative judgements in the explicit condition would be reflected by a BOLD response in the inferior frontal gyrus. In addition, taking into account evidence from visual emotional processing (Critchley et al., 2000; Hariri et al., 2000, 2003), we hypothesized that the amygdala would show a stronger BOLD response in the implicit, and prefrontal and cingulate areas a stronger response in the explicit condition.

Methods

Design

The study followed a two-way ANOVA design with a factor *appraisal level/task instruction* (implicit, explicit) and a factor *emotion* (neutral, anger, fear).

Participants

Sixteen right-handed, healthy volunteers (eight females and eight males; mean age ± standard deviation, 26.0 ± 3.9 years) participated in the study. All participants confirmed that they had no known auditory impairments. Handedness was assessed using the Edinburgh Handedness Inventory (Oldfield, 1971; mean laterality quotient ± standard deviation, 88.9 ± 10.1). Written informed consent was obtained from all participants. The study had been approved by the ethics committee of the canton of Bern.

Stimulus material and tasks

Stimuli were constructed as described previously (Grandjean et al., 2005; Sander et al., 2005). In brief, from the set of Banse and Scherer (1996), eight speakers (four female, four male) were

selected, and from each speaker and each emotion, three excerpts of 750 ms[1] were cut that contained the pseudo-words "*fee gott laish*", "*gosterr*", and "*nyou venzy*". This procedure resulted in 24 stimuli for every emotion category *anger*, *fear*, and *neutral*, summing up to 72 stimuli for the whole experiment. All stimuli were normalized with respect to mean sound pressure level.

In order to ensure comparable recognition of anger and fear prosody, a pilot behavioural experiment was conducted on an independent sample of 20 participants who performed an emotion discrimination task on these stimuli. All individual stimuli were recognized above chance level. Correct identification for individual stimuli ranged from 45% to 100%. Mean performance (±standard deviation) was 74% (±13%) for *anger*, 74% (±12%) for *fear* and 86% (±8%) for *neutral* prosody. In a repeated measures ANOVA on emotion categories, there was a significant main effect of emotion ($F_{2, 38} = 8.3$; $p < .001$). Both emotional categories were not as well recognized as neutral prosody (*anger*: $F_{1, 19} = 7.5$; $p < .05$; *fear*: $F_{1, 19} = 11.0$; $p < .01$), but there was no difference in recognition between *fear* and *anger* ($F_{1, 19} < 1$; *n.s.*). Misattribution patterns differed, with anger and fear misattributed as neutral at a rate of 17%, or 13%, respectively ($|t_{19}| = 2.3$; $p < .05$), and confused at a rate of 9% (*anger* labelled as *fear*), or 13% (*fear* labelled as *anger*) ($|t_{19}| = 2.6$, $p < .05$). Neutral stimuli were misattributed as angry or fearful at a rate of 6%, or 8%, respectively ($|t_{19}| < 1$; *n.s.*). To account for the bias towards responding neutral, an ANOVA was performed on unbiased hit rates (Wagner, 1993) and confirmed the results of the ANOVA on raw hit rates.

In the two functional magnetic resonance imaging (fMRI) tasks, the 72 stimuli were presented in pseudo-random order, including 24 interspersed null events without auditory stimulation. A maximum of three similar events were presented in a row. For both experiments, the same stimuli were used with differing order. During twelve discarded fMRI volumes at the start of each run (36 s), three extra stimuli were presented in order to habituate participants to the task, followed by 18 s of silence. The experiment was programmed in e-prime (Version 1.1.4.4, Psychology Software Tools, Pittsburgh, PA, USA) and run on a personal computer. Responses were collected with a customized two-key device for index and middle finger of the right hand using fiber optics.

In the first task (implicit appraisal), subjects were asked to discriminate the speaker's gender and give a key response with speed instruction (*male*, right index finger, *female*, right middle finger). In the second task (explicit appraisal), subjects were asked to discriminate the emotion of the speaker (*anger*, right index finger, *fear*, right middle finger, *neutral*, double click with right index finger). Instructions were given before each task via headphones and were additionally visible during the task using a mirror/screen system. In both tasks, instruction was given to answer as quickly as possible. The order of the two tasks was kept constant, since we assumed that directly after explicit emotion discrimination, implicit proces-

sing without explicit reference to the emotion category would not be possible.

Imaging

Images were acquired on a 3 T MR-scanner (Trio, Siemens, Erlangen, Germany). Anatomical T_1-weighted volumes were obtained with a three-dimensional magnetization-prepared rapid acquisition gradient echo sequence (MPRAGE) at a voxel size of 1 mm^3 (repetition time *TR*, 1950 ms; echo-time *TE*, 2.15 ms). For fMRI, we used BOLD signal-sensitive T_2^*-weighted gradient-recalled echo-planar imaging (EPI; *TE*, 30 ms; *TR*, 3000 ms; silent interval between consecutive functional volume acquisition, 1000 ms; slice acquisition time, 66 ms). A series of 210 functional whole-brain volumes consisting of 30 contiguous slices 4 mm thick positioned in the intercommissural plane (field of view, 230×230 mm^2; matrix, 64×64 pixels) were acquired. The first twelve volumes were discarded to obtain steady-state longitudinal magnetization and scanner-induced auditory excitation. Between each functional volume acquisition lasting 2000 ms, a silent gap of 1000 ms was inserted (Fig. 1). In every second gap, either a stimulus (prosodic stimulus, 750 ms, see also previous footnote 1) was presented, or this interval was defined as null-event (silence). Within these 1000 ms gaps, stimulus onset delay varied randomly between 50 ms and 200 ms (Fig. 1). We presented 72 stimuli of three categories (*anger*, *fear*, and *neutral*), and interspersed 24 null events in an event-related design via headphones (Commander XG, Resonance Technology, Northridge, CA, USA). To reduce scanner noise, additional earplugs were used, while stimulus volume was accordingly adapted so that stimuli were well perceived.

Data analysis

Data analysis was performed using statistical parametric mapping (SPM 5; Wellcome Trust Centre for Neuroimaging, London, UK; www.fil.ion.ucl.ac.uk/spm) on Matlab (Version 7.1., MathWorks, Natick, MA, USA). For intra-run realignment, the MR-scanner manufacturer's online motion correction was used. Functional images were realigned between runs, corrected for slice timing, normalized to the Montreal Neurological Institute (MNI) T_1-weighted brain template (resampled voxel size: 2×2×2 mm^3), and spatially smoothed (8 mm isotropic full width at half maximum Gaussian kernel) using standard procedures in SPM 5. Each event type was modelled as a separate regressor convolved with a canonical
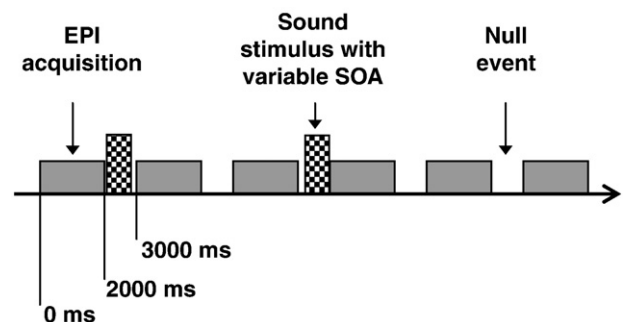
---

[1] A reviewer of this manuscript doubted that it was possible to create speech stimuli of exactly equal length. This is of importance since stimulus duration might influence BOLD responses in the right superior temporal gyrus (Wiethoff et al., 2008). The procedure described here might indeed include silent gaps between syllables or words and therefore cause variability in the effective duration of the stimuli. By analysing intensity contours with three arbitrary intensity thresholds (10, 20, and 30 dB below mean intensity across all sounds), we found a duration of (mean±standard deviation) 703±4.5 ms, 734±2.9 ms, and 744±2.4 ms. Note that the variability in duration is about 100 times smaller than that reported in Wiethoff et al. (2008). There were no significant differences in duration between the three emotion categories.

Fig. 1. Experimental design. *Grey*: Echoplanar image acquisition lasting 2000 ms. *Intervals*: In every second interval (1000 ms), either a stimulus (750 ms) with varying onset (50–200 ms) is presented, or a null-event is interspersed.

hemodynamic response function. In order to account for differences in task difficulty, individual reaction times for each event – which are a more meaningful descriptor of difficulty on a trial-by-trial basis than the actual responses – were modelled as one additional regressor for all event types. In order to account for variance caused by basic stimulus characteristics, the mean fundamental frequency $F0$ of each stimulus was calculated using Praat (Version 4.5, www.praat.org), controlled for octave jumps by visual inspection, and also introduced into the model as one regressor for all event types. These parametric modulators were orthogonalised using a serial Gram–Schmidt algorithm as implemented in SPM 5. In effect, the F0 regressor was orthogonalised with respect to the reaction time regressor, while the mean was removed from the reaction time regressor which was otherwise left unchanged. Both parametric regressors were then convolved with a canonical hemodynamic response function. Due to the correlation of task instruction and task difficulty, the reaction time regressor might obscure task effects. All analysis were therefore repeated without reaction time regressor, confirming all BOLD response clusters with regard to localization and significance levels, and approximating cluster size. Statistical parametric maps were generated from linear contrasts of interest (main effect task instruction, main effect $emotion > neutral$, main effect $anger > fear$, interaction task instruction × emotion) in each participant. A second level random effect analysis (RFX) was then performed using one sample t-tests on contrast images obtained in each participant for each comparison of interest ($df = 15$). We report signal clusters with a voxel-level threshold of $p < 0.001$ (uncorrected), whole-brain corrected for family-wise error (FWE) at cluster level ($p < .05$), using the random field theory approach implemented in SPM 5. Since the amygdala is of special interest in the context of implicit task instruction, we defined a region of interest in the amygdala as described by cytoarchitectonic probability maps (Eickhoff et al., 2007). For clusters extending into this region, small-volume correction for FWE was applied at voxel level ($p < .05$) using a sphere of 30 mm diameter around cluster peak coordinates. In order to evaluate common responses to $anger$ and $fear$ prosody in a more conservative approach, a conjunction analysis ($anger > neutral$) and ($fear > neutral$) was conducted, using a second level two sample t-test and testing against conjunction null (Nichols et al., 2005; see also Friston et al., 2005). Unique responses to these emotion categories were assessed by contrasting $anger$ with $fear$ prosody. For the interaction of task instruction × emotion, regions of interest were defined as brain areas where activity was found in either

main effect. BOLD responses in those regions of interest were small volume corrected for FWE at voxel level ($p < .05$) using a sphere of 30 mm diameter around cluster peak coordinates.

Voxel time courses in responding clusters were averaged for each subject using MarsBaR (Marseille boîte à région d'intérêt; www.marsbar.sourceforge.net). Using Matlab, event-related averaging was performed for each subject and each cluster, and percent signal change was calculated using two pre-stimulus volumes per condition as baseline. Results were then averaged across subjects. We report peak signal change difference between the conditions of interest, and latency of that peak, expressed as interval between stimulus onset and onset of the MR volume during which the peak occurred. Behavioural responses were summarized using R (www.r-project.org) and analysed using the general linear model approach in SPSS (Version 12.0.2, SPSS Inc., Chicago IL, USA). In addition to raw accuracy measures, an unbiased accuracy estimate was calculated for the explicit task, which multiplies the rate of correct responses for a given stimulus category and the rate with which a given response was used correctly (Wagner, 1993).

Results

Behavioural analysis

During the fMRI tasks, accuracy (i.e. correct responses, or hit rate) and reaction times (RT) were collected from each participant. Results are shown in Fig. 2. It should be noted here that due to slightly differing motor difficulty of the required reactions, reaction times have to be interpreted with caution. There was a significant main effect of appraisal level/task instruction in both measures (accuracy: $F_{1, 15} = 121.6$; $p < .001$; RT: $F_{1, 15} = 86.2$; $p < .001$) as well as a significant main effect of emotion (accuracy: $F_{2, 30} = 34.9$; $p < .001$; RT: $F_{2, 30} = 6.3$; $p < .001$). Additionally, the interaction between both factors was significant (accuracy: $F_{2, 30} = 11.3$; $p < .001$; RT: $F_{2, 30} = 3.5$; $p < .05$). A $post\ hoc$ contrast confirmed that accuracy differences between the tasks were more pronounced for $fear$ and $anger$ than for $neutral$ stimuli, while this contrast did not reach significance for reaction times (accuracy: $F_{1, 15} = 15.1$; $p < .01$; RT: $F_{1, 15} = 2.1$; $n.s.$). A $priori$ contrasts were calculated to identify differences between individual emotion categories and are shown in Fig. 2. The difference in performance between anger and fear in the explicit task is almost entirely explained by misattributions of fear as anger (16.4%), while anger was less frequently taken for fear (3.4%; $|t_{15}| = 6.4$;
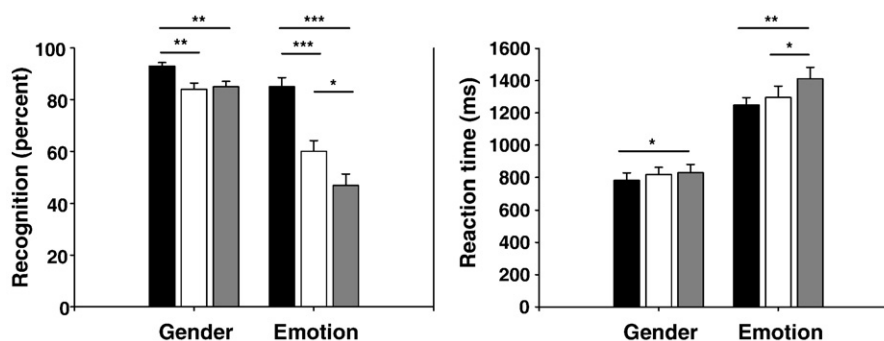

Fig. 2. Behavioural data: response accuracy and reaction times for gender identification (implicit task instruction) and emotion identification (explicit task instruction) of neutral (black bars), anger (open bars), and fear stimuli (grey bars). Results are shown as mean ± standard error. Unbiased hit rates are additionally stated in the text. $Post\ hoc$ contrasts: *$p < .05$, **$p < .01$, ***$p < .001$.

$p<.001$). Anger and fear were misattributed as neutral at a similar frequency (34.4% and 33.9%, respectively, $|t_{15}|<1$). In order to account for such response biases, an unbiased hit rate (Wagner, 1993) was calculated for the explicit task and revealed corrected accuracy of .49, .44, and .40, respectively, for neutral, anger, and fear prosody. In a one-way ANOVA, there was a tendentially significant effect of emotion category ($F_{2, 30}=3.2$; $p=.06$) with fear accuracy differing significantly from neutral accuracy ($F_{2, 30}=7.8$; $p<.05$).

Main effect of appraisal level/task instruction

Brain regions that showed stronger BOLD responses to one of the two tasks are summarized in Table 1 and Fig. 3. During implicit as compared to explicit processing, we observed stronger BOLD responses in the left superior temporal gyrus (STG) and sulcus (STS). Also, a cluster in the right superior and inferior parietal lobule, extending into the postcentral gyrus showed enhanced BOLD response. Two small clusters extending into the amygdala in both hemispheres survived small-volume correction.

During explicit as compared to implicit processing, several clusters showed stronger BOLD signal in the left dorsolateral prefrontal cortex (inferior and middle frontal gyrus [IFG and MFG]), bilateral medial frontal cortex (anterior cingulate [ACC], and supplemental premotor cortex), and bilateral superior parietal lobule and right precuneus. The location of the right parietal areas was posteriomedial of the right hemisphere cluster observed in the opposite contrast, while IFG responses to explicit versus implicit prosody were more superior than responses to emotional versus neutral prosody.

Emotion-specific effects of appraisal level/task instruction

The experimental design using both neutral and emotional prosodic stimuli allowed us to assess task effects specific to emotional as compared to neutral prosody (that is, emotion×task interactions). Such effects were observed for the explicit as compared to the implicit condition. A cluster in the right basal ganglia and another cluster in the right ACC and middle frontal gyrus survived small-volume correction.

Main effect of emotional compared to neutral prosody

Overall effects of emotional as compared to neutral prosody were analysed across tasks and revealed a stronger BOLD signal in the right STG extending into the putamen, in a cluster in the left inferior frontal gyrus and insula, extending into the putamen, and in an extended cluster in the bilateral ACC. In a second step, both unique and common responses to anger and to fear prosody were analysed in a more conservative approach. In order to identify unique responses, both stimulus categories were compared. BOLD responses to anger were more pronounced than to fear in the bilateral basal ganglia, comprising parts of the putamen and caudate head and body. Common responses, that is, brain areas that respond consistently and significantly to both single categories, were identified using a conjunction analysis, testing against conjunction null (Nichols et al., 2005, see also Friston et al.,

Table 1
Main effects of appraisal level/task instruction and of emotional versus neutral prosody on blood oxygenation level-dependent (BOLD) responses

Brain regions	Brodmann area of local maxima	Hemisphere	Voxel number	Voxel Z score	Montreal Neurological Institute brain template coordinates of local maxima	BOLD peak (%)	BOLD peak latency (s)
Implicit > explicit							
Inferior and superior parietal lobule, postcentral gyrus	40 and 7	Right	334	4.64	50, −40, 56; 36, −52, 62; 42, −32, 52	0.05	3.875
Superior temporal gyrus and sulcus	22 and 21	Left	176	3.94	−64, −16, 4; −64, −32, 0	0.05	3.875
Amygdala (small-volume corrected)		Right	8	3.79	30, 2, −14	0.05	3.875
		Left	12	4.39	−32, −10, −20	0.02	6.875
Explicit > implicit							
Superior parietal lobule and precuneus	7 and 19	Bilateral	1593	4.83	−30, −66, 44; −4, −70, 46; −32, −72, 38	0.05	3.875
Precuneus	7 and 31	Right	129	3.71	16, −64, 20; 20, −68, 28; 24,−60, 26	0.03	3.875
Inferior and middle frontal gyrus	45 and 46	Left	201	4.07	−46, 22, 18; −46, 38, 16; −44, 30, 18	0.09	6.875
Superior and middle frontal gyrus, anterior cingulate	6	Bilateral	183	3.93	0, 4, 56; −10, −4, 58	0.07	3.875
Emotion > neutral							
Inferior frontal gyrus, insula, putamen	47	Left	160	4.66	−26, 18, −8; −24, 10, −8; −16, 8, −4	0.03	3.875
Anterior cingulate	32	Bilateral	1486	4.57	8, 14, 42; 6, 28, 34; −4, 14, 44	0.08	3.875
Superior temporal gyrus and putamen	38	Right	129	3.81	30, 16, −10; 52, 16, −8; 24, 12, −6	0.04	3.875
(Anger > neutral) and (fear > neutral)							
Anterior cingulate	24 and 6	Bilateral	1182	4.45	6, 10, 36; −2, 12, 44; −4, 12, 36	0.08	3.875
Middle frontal gyrus	46 and 9	Left	216	3.93	−42, 20, 30; −46, 12, 36	0.09	3.875
Anger > fear							
Caudate and putamen		Bilateral	228	4.54	14, 20, 4; 6, 16, 6; 30, 34, −2	0.07	3.875
Emotion > neutral, explicit > implicit (small volume corrected in regions of interest)							
Putamen		Right	79	4.03	28, 16, 8; 36, 8, 8	0.02	3.875
Anterior cingulate and middle frontal gyrus	6	Right	28	3.88	14, 0, 56	0.02	6.875

All reported clusters survived cluster-level correction at $p<.05$ unless stated otherwise.
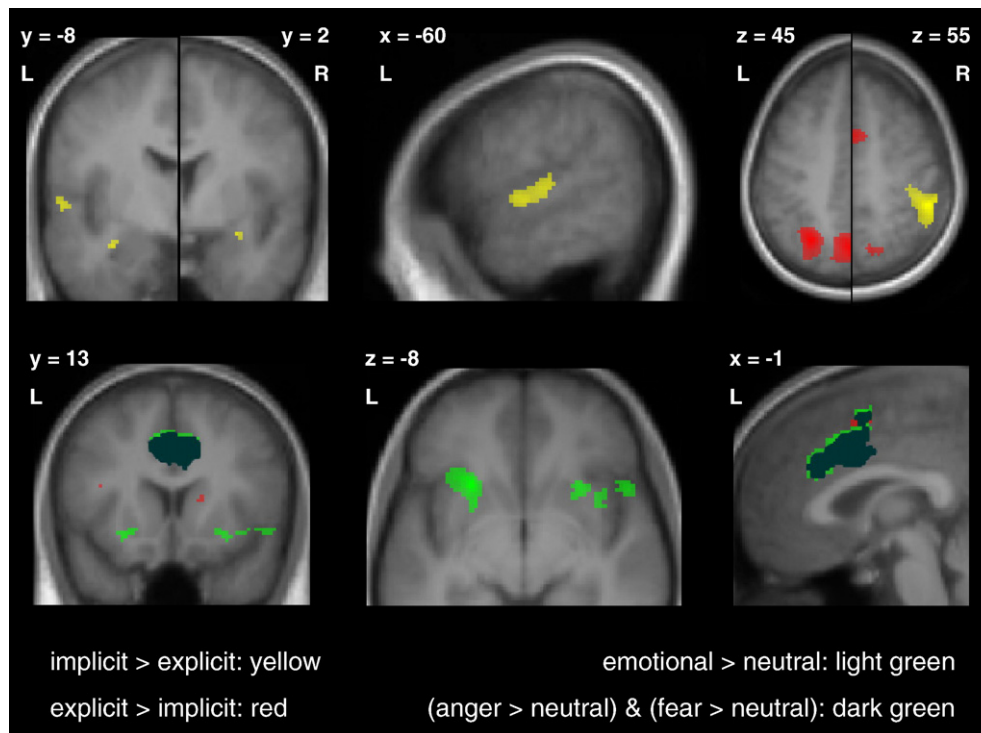
Fig. 3. Blood oxygen level-dependent signal responses to implicit>explicit task instruction (yellow), explicit>implicit task instruction (red), emotional>neutral prosody (green), and conjoint responses to both anger>neutral and fear>neutral prosody (dark green, overlaid on other clusters). Top row left: bilateral amygdala and left superior temporal sulcus (STS); top row middle: left STS; top row right: parietal areas. Bottom row left: left inferior frontal gyrus (IFG), right superior temporal gyrus (STG), bilateral basal ganglia and anterior cingulate (ACC); bottom row middle: left IFG, right STS and bilateral basal ganglia; bottom row right: bilateral ACC.

2005). An extended cluster in the ACC was found that responded consistently to both stimulus classes, as well as a smaller cluster in the middle frontal gyrus.

Discussion

The objective of this study was to investigate and compare implicit and explicit processing of emotional prosody, and to assess common and emotion-specific substrates of appraisal level. In accordance with our hypotheses, we found stronger BOLD responses in the amygdala during implicit and in prefrontal and cingulate areas during explicit processing. Particularly, the inferior frontal gyrus showed stronger responses during explicit processing. This is in line with previous models on explicit evaluation of emotional prosody, while lateralization of that function has been regarded differently (Ethofer et al., 2006a; Schirmer and Kotz, 2006).

Implicit versus explicit processing of prosody

During implicit as compared to explicit processing of prosody, BOLD responses (hence: neuronal activity) were more pronounced in left STS and adjacent areas, and in the right parietal lobe. In addition, small clusters in both amygdalae were activated more during implicit processing.

The left STS has been implied in emotional prosody processing in a paradigm involving an implicit task (Grandjean et al., 2005; Sander et al., 2005) as well with explicit task instructions (Ethofer et al., 2006b). Its role in implicit as opposed to explicit processing has however been reported ambiguously. In two studies investigating phonetic (*implicit*) vs. emotional (*explicit*) judgements on emotional prosody, the left STS has not

been found active under phonetic instructions (Wildgruber et al., 2004, 2005). Another study however reported increased left STG activity in a verbal as opposed to an emotional task, when only *happy* stimuli of the emotional task were analysed, but not when *happy* and *sad* stimuli were accounted for (Buchanan et al., 2000). No contrasts vs. rest are reported in this study. While these results could be interpreted as reduced STG activity during explicit evaluation of *happy* stimuli rather than as increased activity during implicit processing, our results are in favour of a particular function of the left STS during implicit processing of emotional prosody. This left STS activity is not paralleled in studies on implicit processing of visual stimuli (Critchley et al., 2000; Hariri et al., 2000, 2003), which suggests that this function might be specific to auditory emotional stimuli.

Amygdala activity resembles findings from three previous studies investigating appraisal level in the processing of visual emotional stimuli that consistently reported activity in the amygdala region during the more implicit as compared to the more explicit task (Critchley et al., 2000; Hariri et al., 2000, 2003).

Parietal activity in response to implicit prosody processing is consistent with previous reports on parietal activity during phonetic versus emotional processing (Ethofer et al., 2006a; Wildgruber et al., 2005). Task-related activity in this area might be constrained to auditory stimuli, as it has not been reported in implicit processing of visual emotional cues (Critchley et al., 2000; Hariri et al., 2000, 2003).

Explicit versus implicit processing of prosody

During labelling of emotion (explicit appraisal) as compared to labelling of gender (implicit appraisal), we observed

activity in bilateral parietal cortex, including superior parietal lobule and precuneus, in the left dorsolateral prefrontal cortex (DLPFC), including inferior frontal gyrus (IFG), and in a cluster extending into the ACC and the supplemental motor area.

Several studies comparing emotion labelling versus phonetic/semantic labelling of emotional prosody have reported right rather than left IFG activity (Buchanan et al., 2000; George et al., 1996; Wildgruber et al., 2005), and it has been proposed that the right inferior frontal and orbitofrontal cortex is responsible for evaluative judgements of emotional prosody, while the left inferior frontal gyrus serves semantic processing (Schirmer and Kotz, 2006). There is however evidence from a dynamic causal modelling study that both IFGs process prosody in parallel (Ethofer et al., 2006a), and it has been shown that circumscribed left orbitofrontal lesions might impair recognition of emotional prosody (Hornak et al., 2003). Our results might suggest that the left IFG is particularly involved in explicit as opposed to implicit evaluative judgement of affective prosody. However, other possible explanations cannot be ruled out. Attention to the speaker's emotional communication (as required by the explicit task instructions in this study) might enhance efforts to retrieve lexical information for the non-words presented here as compared to attention towards the speakers identity (as required by implicit task instructions). Since the left IFG has been implied in semantic analysis (Schirmer and Kotz, 2006), this might also explain enhanced left IFG responses during explicit processing.

ACC activity has been reported in a study investigating brain responses to explicit processing of visual emotional stimuli (Hariri et al., 2003), although two earlier studies have not reported this activation (Critchley et al., 2000; Hariri et al., 2000). In a study investigating the effect of cognitive load on processing of anxiety, it was found that ACC activity is attenuated by the cognitive load of an unrelated task, thus suggesting that it might reflect higher cognitive appraisal processes that can be diminished by cognitive load (Kalisch et al., 2006). Hence, ACC activity in the explicit condition might be due to task instructions requiring high-level appraisal. On the other hand, the ACC might also serve a specific role in affective prosody processing, particularly in explicit evaluative judgements of affective prosody, since it has been shown that ACC lesions impair recognition of affective prosody (Hornak et al., 2003). The lack of ACC activation in previous studies comparing emotional labelling and phonetic evaluation of emotional prosody can be explained by task instructions involving speech-related instructions in the implicit condition: during phonetic instructions compared to baseline, the ACC was strongly activated (Wildgruber et al., 2004, 2005).

Parietal activity during explicit emotional as compared to phonetic/semantic instructions has been reported in the study of Buchanan et al. (2000), but not in the studies of Wildgruber et al. (2004, 2005). The parietal cortex contains polymodal areas (Bremmer et al., 2001) and might subserve higher analyses of the auditory signal.

Task-related activity in the supplemental motor area might be related to the greater motor difficulty of the explicit task, with three instead of two reaction possibilities, one of which also necessitated a slightly more complicated motor program, that is, a double click.

Effects of emotional versus neutral prosody

During listening to emotional as compared to neutral prosody, we found activity in the right STG, in the left IFG, the bilateral ACC, as well as in the basal ganglia. Emotional versus neutral prosody has been studied previously with both implicit (Grandjean et al., 2005; Sander et al., 2005) and explicit (Kotz et al., 2003) task instructions. The study by Grandjean et al. (2005) is of special interest since it used *angry* stimuli from the same set as our material. In a paradigm similar to the one used in the present study, activity was found in bilateral STS/STG, ACC, postcentral gyrus, and left putamen. With the exception of the left STG and postcentral gyrus, this pattern is closely resembled by our findings, while in the present study, the inferior frontal gyrus was additionally activated. Those differences are mainly attributable to the additional use of *fear* and differing significance thresholds, since when analysing only *anger* vs. *neutral* prosody, we found additional left STG activity at a cluster-level corrected threshold, and at an uncorrected voxel-level threshold comparable to that of Grandjean et al., additional pericentral activity was observed. While our results are therefore in line with a previous study using a similar stimulus set, the contrast emotional vs. neutral prosody cannot be expected to show all areas involved in prosody processing. In the context of emotional prosody stimuli, it might be assumed that also neutral stimuli are evaluated by some brain areas serving emotional prosody processing in general. Greater activity in response to emotional than to neutral prosody might therefore reflect mainly the detection aspect rather than the processing aspect of prosody evaluation. Of special importance here is the ACC that was conjointly and significantly activated both by anger and fear prosody using the conservative conjunction approach. The relevance of the ACC during emotional prosody processing is not yet well understood, although ACC activity has been reported in response to anger vs. neutral prosody (Grandjean et al., 2005), and lesions can impair recognition of emotional prosody (Hornak et al., 2003) which suggests a specific role of this region for detecting emotional prosody. A similar ACC region was activated in the contrast of explicit versus implicit task instructions. This might at first sight suggest that ACC activation during emotional as compared to neutral prosody was in fact due to recruitment of explicit appraisal processes that might be more pronounced when a stimulus is detected as emotional, and thereby potentially relevant. This explanation is however unlikely, given that ACC responses to emotional versus neutral prosody were dramatically more pronounced than during explicit versus implicit processing, again pointing towards a specific ACC function during emotional prosody detection.

It is interesting to note that also in the bilateral basal ganglia, activity was observed during emotional prosody versus neutral prosody. This response was more pronounced for anger than for fear, and it was particularly pronounced during explicit task instructions. Previous studies comparing emotional and neutral prosody under implicit task instructions reported activity in the caudate head (Kotz et al., 2003) and putamen (Grandjean et al., 2005). Also, there is evidence that basal ganglia dysfunction, for example in Parkinson disease, impairs emotional prosody recognition (Pell and Leonard, 2003). Taken together, this points towards a specific function of this brain structure in emotional prosody detection, especially when explicit evaluation is required.

Emotion-specific effects of task instruction

One starting point for the present study was the assumption that appraisal effects might differ for emotional and non-emotional stimuli. Such differences would present as interaction of task instruction×emotion. We observed such effects in the right ACC and basal ganglia, again suggesting a specific function of these areas in the explicit detection of emotional prosody. While previous studies on appraisal levels with comparable paradigms have reported main effects of task instruction and no interaction (Critchley et al., 2000; Hariri et al., 2000, 2003), they were not specifically designed to find such interactions. It should be noted that also in our study, the small number of stimuli per cell of the design does not provide the statistical power that an experiment might achieve that focuses on such interactions. Also, in the present study interactions were specifically analysed in areas where a main effect was present, thus providing greater power for detecting ordinal (but not disordinal) interactions but also increasing the rate of type I errors. The question of how emotion×task interactions are represented in the brain therefore requires further investigation.

Amygdala in emotional prosody processing

Amygdala activity under implicit as opposed to explicit task instructions resembles results from previous studies using visual stimuli (Critchley et al., 2000; Hariri et al., 2000, 2003). An amygdala-cortical network has been proposed where higher-level appraisal processes are represented in prefrontal and cingulate areas and might attenuate amygdala activity related to lower-level processes (Hariri et al., 2000, 2003). It is striking, however, that while emotional faces activate the amygdala more than neutral faces (Whalen et al., 2001), no such effect could be found in our study for emotional versus neutral prosody. Specifically, no effect of fear prosody could be observed in the amygdala, even when only fear vs. neutral prosody was contrasted at a threshold of $p<.001$ uncorrected. Accounts of amygdala function in emotional prosody processing are however ambiguous in the previous literature. Lesion studies suggest that the human amygdala is not as critical for the recognition of affective prosody as it is for recognition of facial affect (Adolphs and Tranel, 1999). This is contrasted by the effects of pure (non-verbal) emotional vocalization (Fecteau et al., 2007; Morris et al., 1999; Phillips et al., 1998; Sander and Scheich, 2001; Seifritz et al., 2003) and simple auditory warning stimuli (Bach et al., 2008), which in most paradigms have shown to elicit strong amygdala responses. This suggests that stimuli with linguistic content might be less susceptible to activate the amygdala than non-linguistic stimuli. An explanation for these findings relies on task difficulty. Although the above cited studies did not require or report stimulus identification performance, it can be speculated that pure emotional vocalization is much easier to recognize than linguistic prosodic cues. Also, visual stimuli used in the above cited studies were very easily recognizable. Therefore, these stimuli are probably more salient and might therefore activate the amygdala more readily, while the more difficult stimuli used in our experiment might require a higher appraisal level *per se*. Note however that interindividual differences in amygdala activation during listening to emotional prosody have been reported, and that these differences might be correlated with individually differing relevance of prosodic stimuli (Schirmer et al., 2008).

Stimulus or response?

All analyses in this article were performed with regard to the presented rather than the identified stimulus category. This is in line with previous studies on appraisal effects on processing of emotional cues. However, when stimulus difficulty is high, this approach might obscure brain responses to the recognition of individual emotion categories, as it does not distinguish between correct and incorrect responses. In addition, error variance might be increased as trials during which participants did not pay attention to stimuli are included in the analysis. One strategy to circumvent these problems would thus be to perform analyses only on correctly identified stimuli. However, the meaning of such an analysis would be rather different between implicit and explicit task instructions. During the implicit task, correct response might be regarded as an indicator of attention, particularly since the task is comparably easy. On the other hand, under explicit instructions and with a comparably difficult task, incorrect responses do not necessarily imply lack of attention. Here, response accuracy is rather an indicator for the perceived emotional meaning. An analysis of correct responses under explicit instructions might therefore reveal brain areas involved in the recognition of an individual emotion category. However, even an analysis limited to explicit task instructions not without problems as stimuli are not necessarily recognized in a dichotomous manner. Uncertainty in recognition might introduce a probabilistic element into responses. Therefore, from a correct response on a given trial it is not possible to infer correct identification of the presented stimulus. For all these reasons, we believe that the presented analysis strategy is best suited to shed light on the effect of appraisal processes. In order to further clarify the recognition of individual emotion categories, analyses according to stimulus identification might however be useful. To overcome the problem of non-dichotomous recognition, one might ask participants to judge the probabilities that a given stimulus belongs to any emotion category.

Limitations

The fixed task order used in this study might suggest that differences between the tasks are in part due to stimulus habituation effects. This possibility cannot be ruled out for the present study. Also, the task used did not control for effects of attention towards stimuli. Another limitation is the difference in difficulty between the tasks. Although main effects of difficulty were controlled by using individual reaction times as a covariate for all analyses, difficulty×task interactions would not be ruled out by this approach. Matched difficulty between tasks would be desirable. Previous studies on appraisal of emotional cues have attempted to provide such paradigms, with mixed results however. While some did not report statistical comparisons of task difficulty (Critchley et al., 2000; Ethofer et al., 2006a,b), others have provided comparable tasks (Buchanan et al., 2000; Hariri et al., 2000; Wildgruber et al., 2004), and two studies report significant differences in difficulty across appraisal levels (Hariri et al., 2003; Wildgruber et al., 2005). Yet another limitation refers to appraisal processes not captured by task instructions. In other words,

although the implicit task was always presented first, it cannot be ruled out that explicit reference to emotion category already took place during the implicit task. This possibility should, however, reduce differences between implicit and explicit tasks and therefore render testing more conservative.

Conclusion

Our study aimed at comparing explicit and implicit processing of affective prosody by avoiding speech-related task instructions for the implicit condition. Our results are in part comparable to those obtained using speech-related task instruction. However, they differ in some important respects. Of note here is amygdala activity during implicit and ACC activity during explicit processing, which has not been observed before during prosody processing but resembles findings from studies using visual cues. Also, inferior frontal gyrus activity during explicit appraisal was observed in the left rather than in the right hemisphere, which might suggest a specific function of this area in explicit evaluative judgements of emotional prosody. The contrast emotional vs. neutral prosody revealed activity in brain areas that have been widely related to emotional prosody processing, that is, the right STS and left inferior frontal gyrus. In addition, we observed activity in the ACC and caudate head. Our study also aimed at investigating emotion-specific task effects. Emotion effects that were more pronounced under explicit task instructions could be shown in the basal ganglia and the ACC. Both areas have been shown to be critical for prosody recognition in lesion studies, but have only rarely been reported in neuroimaging studies comparing emotional and neutral prosodies.

Acknowledgments

References

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

Bach, D.R., Schachinger, H., Neuhoff, J.G., Esposito, F., Salle, F.D., Lehmann, C., Herdener, M., Scheffler, K., Seifritz, E., 2008. Rising sound intensity: an intrinsic warning cue activating the amygdala. Cereb. Cortex 18, 145–150.

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

Bremmer, F., Schlack, A., Shah, N.J., Zafiris, O., Kubischik, M., Hoffmann, K., Zilles, K., Fink, G.R., 2001. Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalences between humans and monkeys. Neuron 29, 287–296.

Buchanan, T.W., Lutz, K., Mirzazade, S., Specht, K., Shah, N.J., 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.

Critchley, H., Daly, E., Phillips, M., Brammer, M., Bullmore, E., Williams, S., Van Amelsvoort, T., Robertson, D., David, A., Murphy, D., 2000. Explicit and implicit neural mechanisms for processing of social information from facial expressions: a functional magnetic resonance imaging study. Hum. Brain Mapp. 9, 93–105.

Dolan, R.J., Morris, J.S., de Gelder, B., 2001. Crossmodal binding of fear in voice and face. Proc. Natl. Acad. Sci. U. S. A. 98, 10006–10010.

Eickhoff, S.B., Paus, T., Caspers, S., Grosbras, M.H., Evans, A.C., Zilles, K., Amunts, K., 2007. Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. Neuroimage 36, 511–521.

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

Ethofer, T., Pourtois, G., Wildgruber, D., 2006c. Investigating audiovisual integration of emotional signals in the human brain. Prog. Brain Res. 156, 345–361.

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

Friston, K.J., Penny, W.D., Glaser, D.E., 2005. Conjunction revisited. Neuroimage 25, 661–667.

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.

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

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

Hariri, A.R., Bookheimer, S.Y., Mazziotta, J.C., 2000. Modulating emotional responses: effects of a neocortical network on the limbic system. Neuroreport 11, 43–48.

Hariri, A.R., Mattay, V.S., Tessitore, A., Fera, F., Weinberger, D.R., 2003. Neocortical odulation of the amygdala response to fearful stimuli. Biol. Psychiatry 53, 494–501.

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.

Johnstone, T., van Reekum, C.M., Oakes, T.R., Davidson, R.J., 2006. The voice of emotion: an FMRI study of neural responses to angry and happy vocal expressions. Soc. Cogn. Affect. Neurosci. 1, 242–249.

Kalisch, R., Wiech, K., Critchley, H.D., Dolan, R.J., 2006. Levels of appraisal: a medial prefrontal role in high-level appraisal of emotional material. Neuroimage 30, 1458–1466.

Killgore, W.D., Yurgelun-Todd, D.A., 2004. Activation of the amygdala and anterior cingulate during nonconscious processing of sad versus happy faces. Neuroimage 21, 1215–1223.

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.

Leventhal, H., Scherer, K.R., 1987. The relationship of emotion to cognition: a functional approach to a semantic controversy. Cogn. Emot. 1, 3–28.

Mitchell, R.L., Elliott, R., Barry, M., Cruttenden, A., Woodruff, P.W., 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 vocalizations. Neuropsychologia 37, 1155–1163.

Nichols, T., Brett, M., Andersson, J., Wager, T., Poline, J.B., 2005. Valid conjunction inference with the minimum statistic. Neuroimage 25, 653–660.

Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9, 97–113.

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.

Phillips, M.L., Young, A.W., Scott, S.K., Calder, A.J., Andrew, C., Williams, 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. B Biol. Sci. 265, 1809–1817.

Pourtois, G., de Gelder, B., Bol, A., Crommelinck, M., 2005. Perception of facial expressions and voices and of their combination in the human brain. Cortex 41, 49–59.

Sander, K., Scheich, H., 2001. Auditory perception of laughing and crying activates human amygdala regardless of attentional state. Brain Res.Cogn Brain Res. 12, 181–198.

Sander, D., Grafman, J., Zalla, T., 2003. The human amygdala: an evolved system for relevance detection. Rev. Neurosci. 14, 303–316.

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

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

Schirmer, A., Escoffier, N., Zysset, S., Koester, D., Striano, T., Friederici, A.D., 2008. When vocal processing gets emotional: on the role of social orientation in relevance detection by the human amygdala. Neuroimage 40, 1402–1410.

Seifritz, E., Esposito, F., Neuhoff, J.G., Luthi, A., Mustovic, H., Dammann, G., von Bardeleben, U., Radue, E.W., Cirillo, G., Tedeschi, G., Di Salle, F., 2003. Differential sex-independent amygdala response to infant crying and laughing in parents versus nonparents. Biol. Psychiatry 54, 1367–1375.

Wagner, H.L., 1993. On measuring performance in category judgment studies of nonverbal behavior. J. Nonverbal Behav. 17, 3–28.

Wambacq, I.J., Shea-Miller, K.J., Abubakr, A., 2004. Non-voluntary and voluntary processing of emotional prosody: an event-related potentials study. Neuroreport 15, 555–559.

Whalen, P.J., Rauch, S.L., Etcoff, N.L., McInerney, S.C., Lee, M.B., Jenike, M.A., 1998. Masked presentation of emotional facial expressions modulate amygdala activation without explicit knowledge. J. Neurosci. 18, 411–418.

Whalen, P.J., Shin, L.M., McInerney, S.C., Fischer, H., Wright, C.I., Rauch, S.L., 2001. A functional MRI study of human amygdala responses to facial expressions of fear versus anger. Emotion 1, 70–83.

Whalen, P.J., Kagan, J., Cook, R.G., Davis, F.C., Kim, H., Polis, S., McLaren, D.G., Somerville, L.H., McLean, A.A., Maxwell, J.S., Johnstone, T., 2004. Human amygdala responsivity to masked fearful eye whites. Science 306, 2061.

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