Blunted feelings: Alexithymia is associated with a diminished neural response to speech prosody

Katharina Sophia Goerlich-Dobre,1,2 Jurriaan Witteman,3,4 Niels O. Schiller,3,4 Vincent J. P. van Heuven,3,4 André Aleman,1,5 and Sander Martens1
1Neuroimaging Center, Department of Neuroscience, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands, 2Department of Psychiatry, Psychotherapy, and Psychosomatics, Medical School, RWTH Aachen University, Aachen, Germany, 3LIBC Leiden Institute for Brain and Cognition, Leiden University, Leiden, The Netherlands, 4LUCL Leiden University Centre for Linguistics, Leiden University, Leiden, The Netherlands, and 5Department of Psychology, University of Groningen, Groningen, The Netherlands

How we perceive emotional signals from our environment depends on our personality. Alexithymia, a personality trait characterized by difficulties in emotion regulation has been linked to aberrant brain activity for visual emotional processing. Whether alexithymia also affects the brain’s perception of emotional speech prosody is currently unknown. We used functional magnetic resonance imaging to investigate the impact of alexithymia on hemodynamic activity of three a priori regions of the prosody network: the superior temporal gyrus (STG), the inferior frontal gyrus and the amygdala. Twenty-two subjects performed an explicit task (emotional prosody categorization) and an implicit task (metrical stress evaluation) on the same prosodic stimuli. Irrespective of task, alexithymia was associated with a blunted response of the right STG and the bilateral amygdalae to angry, surprised and neutral prosody. Individuals with difficulty describing feelings deactivated the left STG and the bilateral amygdalae to a lesser extent in response to angry compared with neutral prosody, suggesting that they perceived angry prosody as relatively more salient than neutral prosody. In conclusion, alexithymia may be associated with a generally blunted neural response to speech prosody. Such restricted prosodic processing may contribute to problems in social communication associated with this personality trait.

Keywords: alexithymia; emotional prosody; amygdala; superior temporal gyrus; inferior frontal gyrus

INTRODUCTION

The perception and interpretation of emotional signals is an important part of social communication. Body gestures and posture, facial expressions as well as the tone of voice provide crucial insight into another person’s mind (Van Kleef, 2009). However, how emotional signals are perceived and interpreted may differ considerably across individuals, and the same emotional signal can evoke a different response in different people (Hamann and Canli, 2004; Ormel et al., 2013). Indeed, recent neuroimaging studies indicate that personality modulates the brain’s response to emotional signals in our environment (e.g. Canli et al., 2001; Hooker et al., 2008; Brück et al., 2011a; Brühl et al., 2011; Frühholz et al., 2011).

Difficulties interpreting emotions lie at the core of alexithymia (‘no words for feelings’), a personality construct referring to a specific deficit in emotional processing (Sifneos, 1973). Individuals scoring high on alexithymia have difficulty identifying, analyzing and verbalizing their feelings, reading emotions from faces (e.g. Parker et al., 2005; Prkachin et al., 2009; Swart et al., 2009) and describing other’s emotional experiences (Bydlowski et al., 2005). These difficulties in the cognitive processing of emotions, which constitute the cognitive alexithymia dimension, may be accompanied by reduced capacities to experience emotional arousal (affective alexithymia dimension). Flattened affect paired with diminished empathy for the feelings of others (Guttman and Laporte, 2002; Gryberg et al., 2010) may lead to a perception of alexithymic individuals as cold and distant (Spitzer et al., 2005) and interpersonally indifferent (Vanheule et al., 2007), resulting in problems in social life.

In the past decade, neuroimaging studies have begun to reveal the neural basis of emotion processing deficits associated with alexithymia. By means of functional magnetic resonance imaging (fMRI), these studies demonstrated aberrant brain activity in individuals scoring high on alexithymia for a variety of emotional processing tasks, (for a meta-analysis, see van der Velde et al., 2013) such as the viewing of facial and bodily expressions of emotions (Berthoz et al., 2002; Mériaux et al., 2006; Kugel et al., 2008; Pouga et al., 2010), during empathy for pain (Bird et al., 2010) and during the imagery of autobiographic emotional events (Mantani et al., 2005). Such differences were not only found during explicit processing (i.e. when participants were asked to explicitly evaluate the emotional dimension of the stimuli) but also during implicit processing (i.e. when participants were asked to direct their attention toward another dimension than emotion or were unaware of the emotional dimension; see De Houwer, 2006), such as during the brief presentation of masked emotional faces (Leweke et al., 2004; Kugel et al., 2008; Duan et al., 2010; Reker et al., 2010; see Gryenberg et al., 2012, for a review).

Previous research into the neural basis of alexithymia has mostly focused on the processing of visual emotional stimuli such as facial or bodily expressions of emotions or emotional pictures and videos. Surprisingly, the impact of alexithymia on the perception of emotional prosody (the melody of speech) has received little attention despite its importance in conveying emotion through the voice in daily conversation. In an electroencephalography (EEG) study using a cross-modal affective priming paradigm, we tested the impact of alexithymia on the N400 component, an indicator of the perception of mismatches in affective meaning, in response to music and speech (Goerlich et al., 2011). Alexithymia correlated negatively with N400 amplitudes for mismatching music and speech, suggesting that people scoring high on this personality trait may be less sensitive to aurally perceived emotions. The suggestion of a reduced sensitivity to emotional speech prosody in alexithymia was confirmed in a further EEG study, in which we additionally observed that alexithymia did not only affect the explicit but also the implicit perception of emotional prosody qualities (Goerlich et al., 2012).
Taken together, these findings suggest that alexithymia is linked to differences in the way the brain processes emotions conveyed through the voice. However, the brain regions underlying such differences remain unknown as neuroimaging studies with the necessary higher spatial resolution to identify them are currently lacking. Therefore, this fMRI study aimed to investigate how alexithymia affects the neural processing of emotional prosody. As the impact of alexithymia may vary depending on whether attention is directed toward the emotional prosodic dimension or not, we investigated both explicit processing (participants categorized emotional prosody, i.e. attention was directed toward the emotional dimension) and implicit processing of the same prosodic stimuli (participants evaluated the metrical stress position, i.e. attention was directed toward a different dimension than emotional prosody).

The neural network underlying emotional prosody perception has been investigated by numerous studies over the past decades (for a recent review, see e.g. Kotz and Paulmann, 2011). A recent meta-analysis of the lesion literature concluded that the right hemisphere is relatively more involved in prosodic processing than the left (Wittteman et al., 2011). Regarding a specific network for prosodic processing, converging evidence suggests the involvement of fronto-temporal regions and subcortical structures (for a recent meta-analysis, see Wittteman et al., 2012). The processing of emotional prosody has been proposed to involve three phases (Schirmer and Kotz, 2006). In the initial phase, basic acoustic properties are extracted, a process presumably mediated by the primary auditory cortex and adjacent middle temporal superior temporal gyrus (STG) (Wildgruber et al., 2009). In the second phase, the extracted acoustic information is integrated into an emotional percept or ‘gestalt’. This process takes place in the superior temporal cortex (STC), with the laterality and anterior–posterior distribution of activity within the STC being sensitive to stimulus-specific features (phonetic medium, valence) as well as to task conditions such as attentional focus (for meta-analyses, see Frühholz and Grandjean, 2013; Wittteman et al., 2012). In the third and final phase, emotional prosody is explicitly evaluated and integrated with other cognitive processes in the inferior frontal gyrus (IFG) (Brück et al., 2011b).

With respect to the involvement of subcortical structures, the amygdala has been implicated by several studies in emotional prosody processing (Sander et al., 2005; Wildgruber et al., 2008; Ethofer et al., 2009; Wiethoff et al., 2009; Leitman et al., 2010; Brück et al., 2011a; Mothes-Lasch et al., 2011; Jacob et al., 2012; Frühholz and Grandjean, in press, but see Adolphs and Tranel, 1999; Adolphs et al., 2001; Wildgruber et al., 2005; Wittteman et al., 2012), with the role ascribed to the amygdala in this context being the initial detection of emotional salience and relevance (Kotz and Paulmann, 2011). The amygdala is a crucial structure for the processing of emotions in general (see Armony, 2013, for a recent review) and has further been suggested to act as a general detector of personal significance (Sander et al., 2003). Amygdala responsiveness seems to depend on personality characteristics, as has recently been demonstrated for the case of neuroticism (Brück et al., 2011a), and several studies reported a reduced responsiveness of the amygdala during the implicit and explicit processing of visual emotional information in alexithymia (Leweke et al., 2004; Kugel et al., 2008; Pouga et al., 2010).

Taken together, previous evidence suggests that the STG, the IFG and the amygdala are involved in the neural processing of emotional prosody. Therefore, these three regions were chosen as a priori regions of interest (ROIs) for the present study. The aim of the study was to investigate the impact of alexithymia on the neural processing of emotional prosody under explicit as well as under implicit task conditions. Given previous evidence of a reduced sensitivity to emotional prosody and a reduced responsiveness of the amygdala to visual emotional information in alexithymia, we hypothesized diminished ROI activity during the implicit and explicit processing of emotional prosody with higher levels of alexithymia.

**METHODS**

**Participants**

A total of 22 subjects (nine males; mean age 24.8 ± 5.3) participated in the study. Three subjects performed at chance level on the implicit task while performing normally on the explicit task. Therefore, analyses of the implicit task did not consider the data of these subjects. All participants were right-handed as determined by the Edinburgh Handedness Inventory (mean 88.11 ± 11.32, minimum 67). Participants were native speakers of Dutch, had normal or corrected-to-normal vision, no hearing problems and no psychiatric or neurological disorder in present or past. All participants gave informed consent prior to the experiment and received compensation for participation. The study was approved by the local ethics committee and conducted in accordance with the Declaration of Helsinki.

**Materials**

Pseudowords (see Appendix A) with a bisyllabic structure were generated. All pseudowords obeyed Dutch phonotactics and were verified for absence of semantic content. All pseudowords were expressed with neutral, (pleasantly) surprised and angry prosody, with stress on the first and second syllable by two professional actors (one male and one female). Stimuli were recorded at 16 bit resolution and a 44.1 kHz bit sample rate in a sound proofed cabin. Surprised and angry in addition to neutral prosody was chosen to sample positive and negative emotions that are both considered to be approach emotions (thus, there was no confounding role of the approach-withdrawal dimension). Items were intensity normalized (i.e. did not differ in mean intensity between categories) and had a mean duration of 756 ± 65 ms. In line with previous literature (Scherer, 2003), the emotional categories differed from neutral prosody in mean F0 and F0 variability, and anger additionally in intensity variability (Table 1). The validity of the intended prosodic contrasts was verified by a panel of five healthy volunteers who classified each stimulus in a forced choice task (categories: angry, surprised, happy, sad, neutral and other). Only pseudowords whose prosodic category was classified correctly by at least 80% of the participants for both actors were chosen.

Two categorization tasks were created presenting identical stimuli in an implicit and an explicit emotional task. For the implicit task, the subjects’ attention was directed toward a non-emotional dimension as they were asked to decide whether the metrical stress lay on the first or the second syllable of the pseudoword. For the explicit task, the subjects’ attention was explicitly directed toward the emotional prosody dimension as subjects decided whether the pseudowords were spoken with neutral, angry or surprised intonation. Tasks and instructions were identical except for the words that instructed participants to either respond to non-emotional (implicit task) or emotional (explicit task) prosodic characteristics.

From the pool of validated stimuli, 32 items of each emotional category were selected, half of which had metrical stress on the first syllable, the other half on the second syllable. Speaker gender was balanced across items.

**Procedure**

Each subject completed both tasks (12 min each). The implicit task was always presented first to prevent subjects from devoting attention explicitly to the emotional dimension of the stimuli in this task. Subjects were instructed that they would hear a nonsense word and to categorize the task-relevant category (emotion or metrical stress) as fast and
accurately as possible with a right-hand button press. Assignment of individual categories to response buttons was counterbalanced across subjects. Participants were instructed that they could respond while the stimulus was still playing (i.e. RT was recorded from the onset of the stimulus).

Subjects first practiced the tasks in the scanner with simulated scanner noise until reaching at least 75% performance accuracy. Then, the experimental trials started which encompassed a total of 96 trials (32 items per emotional category). Throughout the experiment, a black fixation cross was presented in the center of a gray background. Auditory stimuli were presented binaurally through MR-compatible headphones and a trial ended 2 s after stimulus onset. Stimuli were presented in an event-related fashion with a jittered inter-stimulus interval (between 4 and 8 s). The order of stimulus presentation was pseudo-random with the restriction of no more than two consecutive presentations of the same stimulus category. Subjects were instructed to fix their gaze on the fixation cross throughout the experiment. Stimulus presentation was controlled using E-prime 1.2, and stimulus material was presented at 16 bit resolution and a 44.1 kHz sampling frequency at a comfortable intensity level. All subjects reported that the stimulus could be perceived clearly despite the scanner noise.

### Twenty-item Toronto Alexithymia Scale

The 20-item Toronto Alexithymia Scale (TAS-20) is the most widely used measure of alexithymia with a demonstrated validity, reliability and stability (Bagby et al., 1994a,b). A validated Dutch translation of the scale was used (Kooiman et al., 2002). The TAS-20 consists of 20 self-report items rated on a five-point Likert scale (1: strongly disagree, 5: strongly agree), with five negatively keyed items. It comprises three subscales: (1) difficulty identifying feelings (e.g. ‘I often don’t know why I’m angry’), (2) difficulty describing feelings (e.g. ‘I find it hard to describe how I feel about people’) and (3) externally oriented thinking (e.g. ‘I prefer talking to people about their daily activities rather than their feelings’). Possible scores range from 20 to 100, higher scores indicate higher degrees of alexithymia.

### State-Trait Anxiety Inventory

The State-Trait Anxiety Inventory (STAI) (Spielberger et al., 1970) differentiates between the temporary condition of ‘state anxiety’ and the more general and long-standing quality of ‘trait anxiety’. For this study, the Trait Anxiety version of the STAI (T-STAI) was used to control for trait anxiety, which has been reported to be closely linked to alexithymia (Berthoz et al., 1999). The T-STAI evaluates relatively stable aspects of anxiety proneness (general states of calmness, confidence and security), and thus refers to a general tendency to respond with anxiety to perceived threats in the environment. The scale consists of 20 items rated from 1 (almost never) to 4 (almost always); higher scores indicate more trait anxiety.

### fMRI data acquisition

Imaging data were acquired on a Philips 3.0 T Achieva MRI scanner with an eight-channel SENSE head coil for radiofrequency transmission and reception (Philips Medical Systems, Best, The Netherlands). For each task, whole-brain fMRI data were acquired using T*2-weighted gradient-echo echo-planar imaging (EPI) with the following scan parameters: 355 volumes (the first five were discarded to reach signal equilibrium); 38 axial slices scanned in ascending order; repetition time (TR) = 2200 ms; echo time (TE) = 30 ms; flip angle = 80°; FOV = 220 × 220 mm; 2.75 mm isotropic voxels with a 25 mm slice gap. A high-resolution anatomical image (T1-weighted ultra-fast gradient-echo acquisition; TR = 9.75 ms; TE = 4.59 ms; flip angle = 8°; 140 axial slices; FOV = 224 × 224 mm; in-plane resolution 0.875 × 0.875 mm; slice thickness = 1.2 mm) and a high-resolution T*2-weighted gradient-echo EPI scan (TR = 2.2 s; TE = 30 ms; flip angle = 80°; 84 axial slices; FOV = 220 × 220 mm; in-plane resolution 1.96 × 1.96 mm; slice thickness = 2 mm) were additionally acquired for registration to standard space.

### fMRI data preprocessing

Prior to analysis, all fMRI data sets were submitted to a visual quality control check to ensure that no gross artifacts were present in the data. Data were analyzed using FMRIB Software Library (FSL) (www.fmrrib.ax.ac.uk/fsl), version 4.1.3. The following preprocessing steps were applied to the EPI data sets: motion correction, removal of non-brain tissue, spatial smoothing using a Gaussian kernel of 8 mm full width at half maximum, grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor and a high-pass temporal filter of 70 s (i.e. $\geq 0.07$ Hz). The dataset was registered to the high-resolution EPI image, the high-resolution EPI image to the T1-weighted image and the T1-weighted image to the 2 mm isotropic MNI-152 standard space image (T1-weighted standard brain averaged over 152 subjects; Montreal Neurological Institute, Montreal, QC, Canada). The resulting transformation matrices were then combined to obtain a native to MNI space transformation matrix and its inverse (MNI to native space).

### Data analysis

#### Behavioral data analysis

Accuracy (ACC, proportion correct responses) and reaction times (RT for correct responses) were entered as dependent variables in analyses of covariance (ANCOVA) with the within-subjects factors Task (Explicit vs Implicit), Emotion (Neutral vs Angry vs Surprised), Sex as a between-subjects factor, and scores on the TAS-20 alexithymia questionnaire and the T-STAI anxiety scale included as covariates.

#### fMRI data analysis

Data analysis was performed using FEAT (FMRI Expert Analysis Tool) Version 5.98, part of FSL. In native space, the fMRI time series was analyzed in an event-related fashion using the General Linear Model with local autocorrelation correction applied (Woolrich et al., 2004).

---

### Table 1: Acoustic properties of the prosodic stimuli per emotional category

<table>
<thead>
<tr>
<th></th>
<th>Neutral</th>
<th>Anger</th>
<th>Surprise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean intensity (dB)</td>
<td>79.45</td>
<td>79.27</td>
<td>80.64</td>
</tr>
<tr>
<td>Mean variation (s.d.) intensity</td>
<td>8.84</td>
<td>10.74</td>
<td>8.83</td>
</tr>
<tr>
<td>Mean f0 (Hz)</td>
<td>180.73</td>
<td>281.35</td>
<td>282.46</td>
</tr>
<tr>
<td>Mean variation (s.d.) f0</td>
<td>44.78</td>
<td>78.56</td>
<td>101.19</td>
</tr>
<tr>
<td>Mean total duration (s)</td>
<td>0.79</td>
<td>0.76</td>
<td>0.72</td>
</tr>
</tbody>
</table>

### Table 2: Results of the whole-brain cluster-corrected analysis for emotional vs neutral prosody across tasks

<table>
<thead>
<tr>
<th>Area</th>
<th>Cluster size (voxels)</th>
<th>$P_{	ext{FDR}}$ value</th>
<th>$Z$ score</th>
<th>Cluster peaks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right STG</td>
<td>673</td>
<td>0.05</td>
<td>4.48</td>
<td>60 - 10 - 6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.95</td>
<td>60 - 10</td>
<td>60 - 10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.95</td>
<td>64 - 4</td>
<td>60 - 10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.71</td>
<td>60 - 30</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.47</td>
<td>72 - 28</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.42</td>
<td>52 - 42</td>
<td>8</td>
</tr>
</tbody>
</table>
For both runs, onset of each of the stimulus categories (neutral, angry and surprised) was modeled separately as an event with an 800 ms duration at the first level. Each effect was modeled on a trial-by-trial basis using a square wave function convolved with a canonical hemodynamic response function and its temporal derivative. At second level, a whole-brain analysis was performed across both tasks to examine the main effect of emotional (angry and surprised) vs neutral prosody, using clusters determined by \( P < 0.01 \) \((z > 2.3)\), and a cluster-corrected significance threshold \( P_{\text{FDR}} < 0.05 \). If errors were present, these trials were included in the model but not in the contrasts of interest.

**ROI analyses**

ROI analyses were performed on three a priori regions: the STG (mid and posterior part), the IFG (pars opercularis) and the amygdala (all bilaterally). Anatomical ROIs for these regions were created as defined by the Harvard-Oxford cortical structural atlas (http://www.fmrib.ox.ac.uk/fsl/data/atlas-descriptions.html#ho). Mean \( z \)-scores of each ROI were calculated for each stimulus category (anger, surprise and neutral) against baseline for each task and extracted for each participant using Featquery (http://www.fmrib.ox.ac.uk/fsl/feat5/featquery.html). These \( z \)-transformed parameter estimates indicate how well the mean signal of each ROI is explained by the model. To identify emotion-specific ROI activity, mean \( z \)-scores were further extracted for the contrasts angry vs neutral and surprised vs neutral prosody.

**Statistical analyses of ROI data**

The first analysis used the \( z \)-scores against baseline as a dependent variable in a repeated measures (RM) ANCOVA with Task (Explicit vs Implicit), ROI (STG vs IFG vs Amygdala), Hemisphere (Left vs Right) and Emotion (Angry vs Surprised vs Neutral) as within-subject factors, Sex as between-subject factor and TAS-20 alexithymia and T-STAI anxiety scores as covariates. This analysis served to identify the impact of alexithymia on ROI activity in response to angry, surprised and neutral prosody compared to baseline (scanner noise).

In the second analysis, emotional prosody (angry and surprised) was contrasted to neutral prosody in an RM-ANCOVA using the mean \( z \)-scores of the contrasts angry vs neutral and surprised vs neutral prosody as dependent variable, with all other factors being identical to the first analysis. In keeping with the common procedure of contrasting emotional to neutral stimuli employed in the alexithymia literature, this analysis served to identify the impact of alexithymia on ROI activity specifically in response to emotional (angry/surprised) relative to neutral prosody.

For both analyses, follow-up tests and Pearson’s correlations were conducted to identify the sources of the observed effects. As the aim of this study was to investigate the influence of alexithymia on prosodic processing, results will be reported with a focus on main effects of and interactions with the TAS-20 alexithymia scale (the effects of Task and Emotion will be reported elsewhere).
RESULTS
Behavioral data
TAS-20 alexithymia scores ranged from 26 to 60 (mean: 42.18 ± 7.45), i.e. none of the participants had clinical alexithymia (cut-off score ≥ 61, see Taylor et al., 1997). Alexithymia scores were unrelated to age (r = −0.143, P = 0.526) and did not differ between male and female participants (t = 1.19, P = 0.245). TAS-20 scores were significantly correlated with T-STAI anxiety scores (r = 0.498, P = 0.018), indicating more trait anxiety in individuals with higher alexithymia scores. T-STAI scores were included in all analyses as covariates of no interest to control for the impact of trait anxiety.

Analysis of the accuracy data revealed no significant effect of Task [F(1, 15) = 2.695, P = 0.121]. Neither TAS-20 nor T-STAI scores showed a significant effect on accuracy [TAS-20: F(1, 15) = 1.290, P = 0.274, STAI: F < 1], suggesting that alexithymia did not impair the identification of neutral, angry, and surprised prosody (explicit

Fig. 2 ROI activity for emotional > neutral prosody across tasks at the threshold P_{uncorr} < 0.001.

Fig. 3 Left-hemispheric (left) and right-hemispheric (right) ROI activity (mean z-scores) in response to angry, surprised and neutral prosody compared with baseline, in the implicit (top) and the explicit prosody task (bottom). Error bars indicate the standard error of the mean.
task) and of metrical stress (implicit task). Analysis of RT data likewise revealed no significant main effects or interactions.

fMRI data

Whole-brain analysis

The whole-brain cluster-corrected \((P_{FDR} < 0.05)\) analysis for emotional > neutral prosody across both tasks revealed a large cluster in the right STG (Table 2). Figure 1 shows the global brain activity for prosodic stimuli compared with baseline (scanner noise) across tasks and Figure 2 shows the specific brain activity for emotional compared to neutral prosody across tasks.

ROI analyses

Prosody vs baseline. Figure 3 shows the mean ROI activity (STG, IFG and amygdala) in the left and right hemisphere in response to angry, surprised and neutral prosody vs baseline in the implicit (top panel) and the explicit (bottom panel) prosody task. RM-ANCOVA with z-scores of angry, surprised and neutral prosody vs baseline as dependent variable revealed a significant main effect of TAS-20 alexithymia scores as a between-subjects factor \([F(1,15) = 7.119, P = 0.018]\), T-STAI scores and Sex as between-subjects factors were not significant \((F < 1)\). There was no main effect of Emotion and no interaction of Emotion with alexithymia, suggesting that effects applied to neutral prosody as well as to angry and surprised prosody. There was a significant main effect of ROI \([F(2,30) = 3.663, P = 0.038]\) and Hemisphere \([F(1,15) = 4.785, P = 0.045]\) as well as a three-way interaction ROI × Hemisphere × TAS-20 scores \([F(2,30) = 3.698, P = 0.037]\).

Follow-up tests on the factor ROI showed a significant main effect of TAS-20 scores as a between-subjects factor for the amygdala \([F(1,15) = 8.302, P = 0.011]\), suggesting reduced activity of the left and right amygdala for neutral, angry and surprised prosody in individuals with high alexithymia scores. For the STG, there was a trend for TAS-20 alexithymia scores as a between-subjects factor \([F(1,15) = 3.883, P = 0.068]\). Furthermore, a main effect of Hemisphere \([F(1,15) = 8.353, P = 0.011]\) and a Hemisphere × TAS-20 scores interaction \([F(1,15) = 8.740, P = 0.010]\) was found, indicative of reduced activity of the right STG for neutral, angry and surprised prosody with increasing scores on alexithymia. Pearson’s correlations confirmed these effects. Figure 4 visualizes the correlations of alexithymia with activity of the left and right amygdala and the right STG for neutral, angry and surprised prosody vs baseline. Additional Pearson’s correlations on the three alexithymia facets showed that none of these correlations was driven by a particular facet but rather by the entire personality construct. For the IFG, there was no significant main effect or interaction with TAS-20 alexithymia scores.

Emotional vs neutral prosody. A second RM-ANCOVA with z-scores of angry vs neutral and surprised vs neutral prosody as dependent variable was conducted to identify emotion-specific effects of alexithymia. The results revealed a trend toward a main effect
of TAS-20 alexithymia scores as a between-subjects factor \( [F(1,15) = 4.027, P = 0.063] \), and no main effect of T-STAI scores and Sex \( (F < 1) \). There was a significant Task \( \times \) ROI \( \times \) Hemisphere \( \times \) Emotion \( \times \) TAS-20 interaction \( [F(2,30) = 7.989, P = 0.002] \). Follow-up Pearson’s correlations indicated positive correlations of alexithymia with activity of the left amygdala during the explicit evaluation of angry (>neutral) prosody \( (r = 0.547, P = 0.015) \). During the implicit perception of angry (>neutral) prosody, alexithymia correlated significantly with activity of the left \( (r = 0.582, P = 0.009) \) and right amygdala \( (r = 0.515, P = 0.024) \), and with activity of the left STG \( (r = 0.481, P = 0.037) \). A comparable pattern was observed in the right STG; however, this correlation was only marginally significant \( (r = 0.403, P = 0.087) \). Figure 5 shows the significant correlations of alexithymia with the bilateral amygdalae and the left STG for angry vs neutral prosody.

However, rather than indicating higher ROI activity with increasing alexithymia scores, these seemingly positive correlations were caused by a relatively stronger ROI deactivation in response to neutral prosody (>baseline) than to angry prosody (>baseline) with increasing alexithymia scores, as visualized in Figure 6. Thus, individuals with higher alexithymia scores deactivated the left STG and the bilateral amygdalae less in response to angry prosody than in response to neutral prosody, suggesting that angry prosody was perceived as relatively more salient than neutral prosody.

Additional Pearson’s correlations with the three alexithymia facets showed that the correlation with the left amygdala in the explicit and implicit task and the correlation with the left STG in the implicit task were mainly driven by the facet ‘difficulty describing feelings’.

**DISCUSSION**

The aim of this study was to investigate the impact of alexithymia on the neural processing of emotions conveyed by speech prosody. Our initial analysis contrasting prosody to baseline showed that activity of the right STG and the bilateral amygdalae for angry, surprised and neutral prosody was reduced with increasing scores on alexithymia, both during implicit and explicit emotional processing. When specifically contrasting emotional vs neutral prosody, we observed a relatively stronger deactivation of the left STG and the bilateral amygdalae for neutral compared with angry prosody, particularly in individuals with difficulty describing their feelings.

This study is the first to investigate the neural basis of emotional prosody processing in alexithymia. Our finding of modulated brain activity in response to emotional prosody confirms previous findings of alexithymia-related deficits in visual emotional processing (e.g. Berthoz et al., 2002; Kano et al., 2003; Karlsson et al., 2008; Pouga et al., 2010) and suggests that such aberrant emotional processing extends to the auditory level. Our results further confirm previous EEG findings indicating a reduced sensitivity to the emotional qualities of
speech prosody in alexithymia (Goerlich et al., 2011, 2012), and extend these findings by suggesting that such blunted processing may be localized to the STG and the amygdalae. In line with our result of a deactivation in these regions, the same regions have previously been observed to be less responsive to facial expressions of emotion in alexithymia (for a review, see Grynberg et al., 2012). Reduced STG activity has been found during the implicit (masked) processing of surprised (Duan et al., 2010) and happy and sad faces (Reker et al., 2010). Deactivation of the amygdala has been observed in response to fearful bodily expressions (Pouga et al., 2010) and during the implicit and explicit processing of emotional faces (Leweke et al., 2004; Kugel et al., 2008). The present findings extend these results by suggesting that the amygdalae and the STG also show a blunted response to emotions conveyed through the voice with increasing alexithymia scores. In addition, they confirm that also in the auditory domain, alexithymia affects emotional processing not only when attention is explicitly directed toward the emotional dimension but also when emotion is implicitly processed.

In favor of this interpretation, alexithymia was associated with reduced activity in the STG and the amygdala, regions that are involved in earlier phases of emotional prosody processing (Wildgruber et al., 2009; Brück et al., 2011b; Kotz and Paulmann, 2011), but not in the IFG. The IFG is thought to mediate the final phase of emotional prosody processing, in which emotional prosodic information is explicitly evaluated and integrated with other cognitive processes (Brück et al., 2011b). The lack of a modulation of IFG activity in relation to alexithymia could be due to the good performance of participants and the absence of behavioral differences as a function of alexithymia. Taken together, the here observed pattern of alexithymia-related differences in behavioral performance for angry and sad prosody either, despite of significant differences at the electrophysiological level. In addition, a recent study investigating prosodic processing as a function of neuroticism found a similar pattern of differences at the neural level in the absence of behavioral differences (Brück et al., 2011a). Thus, it appears that deficits in emotional prosody processing in a non-clinical alexithymia sample such as the current one might be of a rather subtle nature, and can thus be detected at the neural level even though they do not tend to surface at the behavioral level. A subtle deficit in emotional prosody processing seems not surprising considering that alexithymic individuals are generally high-functioning and socially adapted people with a pronounced tendency to social conformity.

In favor of this interpretation, alexithymia was associated with reduced activity in the STG and the amygdala, regions that are involved in earlier phases of emotional prosody processing (Wildgruber et al., 2009; Brück et al., 2011b; Kotz and Paulmann, 2011), but not in the IFG. The IFG is thought to mediate the final phase of emotional prosody processing, in which emotional prosodic information is explicitly evaluated and integrated with other cognitive processes (Brück et al., 2011b). The lack of a modulation of IFG activity in relation to alexithymia could be due to the good performance of participants and the absence of behavioral differences as a function of alexithymia. Taken together, the here observed pattern of alexithymia-related modulations of neural activity in areas involved in earlier phases of emotional prosody perception on the one hand, and the...
absence of neural modulations at the final stage on the other hand suggests that alexithymia might predominantly affect earlier stages of prosodic processing. These processing differences may be compensated in the final (pre-response execution) stage of this process, resulting in adequate behavioral performance.

Interestingly, we found that individuals with high alexithymia scores showed reduced ROI activity not only for emotional (angry and surprised) but also for neutral prosody, irrespective of whether they focused on emotional aspects of the stimuli or not. In fact, direct comparisons between angry and neutral prosody revealed that alexithymia was associated with a significantly higher deactivation of the left STG and the bilateral amygdalae for neutral compared with angry prosody. This association was driven particularly by the alexithymia facet difficulty describing feelings. On the one hand, this suggests that also individuals with difficulty describing feelings perceive an angry tone of voice as relatively more salient and attention-capturing than a neutral one, in line with the well-known phenomenon of emotional stimuli being more attention-capturing than neutral ones (for a review on emotional attention, see Vuilleumier, 2005). On the other hand, this finding indicates that individuals with difficulty describing their feelings may assign less personal significance to human voices conveying emotions, and even less to those using neutral prosody. This would indicate an opposite pattern than in psychopathological conditions such as borderline personality disorder, in which neutral faces can be perceived as threatening (Wagner and Linehan, 1999) and elicit the same degree of amygdala hyperactivity as emotional faces (Donegan et al., 2003), and in schizophrenia, in which amygdala deactivations for fearful compared with neutral faces in fact resulted from increased amygdala activity for neutral faces (Hall et al., 2008). Recent advances in neuroscience suggest that the role of the amygdala does not seem to be restricted to emotional stimuli but that this structure may represent a more general relevance detector for salient, personally and socially relevant, or novel stimuli (for reviews, see Armony, 2013; Jacobs et al., 2012; Sander et al., 2003). For instance, amygdala activity was found to be higher for neutral stimuli if these were socially relevant compared with neutral, non-social stimuli (Vrtička et al., 2012). In this line of reasoning, our findings of diminished neutral responses to human voices regardless of emotionality could hint to the existence of a more general deficit in the processing of not only emotional but generally socially relevant information including speech prosody in alexithymia. However, this hypothesis should be considered speculative and remains to be tested in future studies.

Limitations

While the use of the TAS-20 scale facilitates comparability of our results to previous findings, it should be noted that this scale assesses only the cognitive dimension of alexithymia. Recent evidence suggests that the affective dimension of alexithymia, the dimension of emotional experience, may differentially affect emotional processing (Moormann et al., 2008; Bermond et al., 2011). In addition, while this study controlled for the influence of trait anxiety, alexithymia may also be associated with depression (Picardi et al., 2011) not assessed here, and levels of depression may alter the perception of emotional prosody (Naranjo et al., 2011; for a review, see Garrido-Vásquez et al., 2011). Thus, it may be worthwhile to take both alexithymia dimensions into account and to additionally control for levels of depression in future studies on prosodic perception in alexithymia.

Conclusions

Alexithymia seems to be associated with a blunted response of the STG and the amygdalae to speech prosody. This diminished response does not seem to be specific for emotional prosody but occurs also for neutral prosody, hinting to the possibility of a more general deficit in the processing of socially relevant information in alexithymia. Neural alterations in the processing of speech prosody may contribute to problems in social communication associated with this personality trait.

REFERENCES


