

# Cerebral processing of emotional prosody: a network model based on fMRI studies

Dirk Wildgruber<sup>1</sup>, Thomas Ethofer<sup>1</sup>, Benjamin Kreifelts<sup>1</sup> and Didier Grandjean<sup>2</sup>

<sup>1</sup>Department of Psychiatry and Psychotherapy, University of Tübingen, Germany

<sup>2</sup>Swiss Center for Affective Sciences, University of Geneva, Switzerland

dirk.wildgruber@med.uni-tuebingen.de

## Abstract

Comprehension of information conveyed by emotional tone of speech is highly important for successful social interactions. Regarding the underlying neurobiological mechanisms, successive steps of cerebral processing involving auditory analysis within the temporal cortex and evaluative judgements within the frontal lobes have been differentiated (Schirmer & Kotz, 2006; Wildgruber et al., 2006). To further disentangle the impact of stimulus properties and appraisal levels a series of fMRI studies has been performed. The results of these studies indicate a strong association of cerebral responses and acoustic properties of the stimuli in some regions (stimulus-driven effects), whereas other areas showed modulation of activation linked to the focussing of attention to specific task components (task-dependent effects). Based on these findings a refined model of prosody processing and cross-modal integration of emotional signals from face and voice is postulated.

## 1. Stimulus-driven effects

Enhanced activation within the middle section of the superior temporal cortex (mid-STC) has been observed in response to human voices as compared to other acoustic signals (Belin, 2000). Regarding emotional prosody, Grandjean and collaborators (Grandjean, 2005) have demonstrated increasing responses within this region related to anger prosody independent of the participant's spatial attention during a dichotic listening task (Fig. 1).

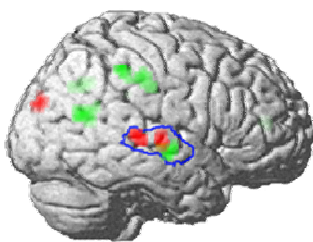


Figure 1: Activations related to angry prosody independent of spatial attention (red) and responses linked to spatial attention (green) rendered on the cortex of a standard brain. Activation within the mid-STC is encircled with blue line (adapted from Grandjean et al., 2005).

In another fMRI study emotional adjectives spoken in happy, angry or neutral intonation were presented to healthy subjects. At semantic level, stimuli comprised unpleasant, pleasant and neutral words. Emotions conveyed by word content and prosody were varied independently. In separate runs,

participants either rated the emotional valence of verbal content or the emotional valence of prosody. Independent of the task, enhanced activation within the mid-STC was associated to increasing intensity of emotional prosody (Ethofer et al., 2006a, Fig. 2).

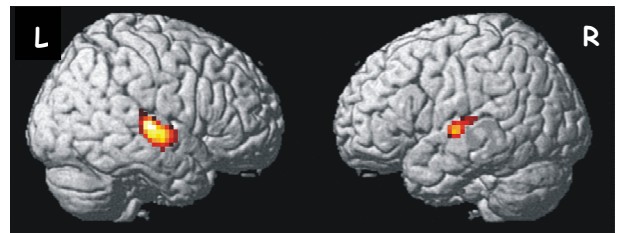


Figure 2: Stimulus-driven effects. Regions characterized by a significant linear relationship between hemodynamic responses and prosodic emotional intensity independent of the task (evaluation of word content or prosody, adapted from Ethofer et al., 2006a).

In a further study passive listening to adjectives and substantives with neutral word content, spoken in five different emotional intonations (happy, neutral, fearful, angry and alluring) was evaluated. All four emotional categories induced stronger responses within the right mid-STC than neutral stimuli (Fig. 3). These responses were significantly correlated with several acoustic parameters (stimulus duration, mean intensity, mean pitch and pitch variability).

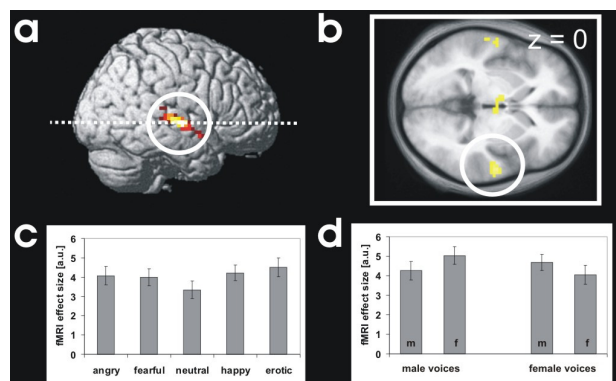


Figure 3: Activation during passive listening to emotional prosody (emotional prosody > neutral prosody). (a) Projection of significant activation rendered on the right hemisphere of a standard brain. (b) Transversal slice showing activation of mid-STC and thalamus. (c) Effect size in right mid-STC for the different prosodic categories. (d) Gender dependent responses during perception of alluring prosody (m = male listener, f = female listener). Error bars represent standard errors of the mean.

A multiple regression analysis revealed that none of the parameters alone could explain the observed activation pattern. The conjoint effect of acoustic parameters, however, sufficiently explained increase of responses within the right mid-STC. Therefore, an important contribution of this area to the integration of several acoustic parameters into an emotional percept has been suggested (Wiethoff et al., 2008). Moreover, an analysis of interaction effects between gender of the speaker and the listener revealed a cross-gender interaction with increasing responses to the voice of the opposite sex in male and female subjects. This effect was confined to an alluring tone of speech in behavioural data as well as hemodynamic responses within the mid-STC (Fig. 3d). The response pattern of the mid-STC, thus, indicates a particular sensitivity to emotional voices with a high behavioural relevance for the listener (Ethofer et al., 2007).

## 2. Task-dependent effects

If attention of the participants is explicitly directed towards the emotional auditory stimulus by specific task instructions further cerebral regions are activated. When subjects explicitly judged the emotional valence conveyed by prosody, for example, increased responses within the right posterior STC (post-STC) and the bilateral inferior frontal cortex were observed as compared to explicit evaluation of the verbal content of identical stimuli (Ethofer et al., 2006b, Fig. 4a).

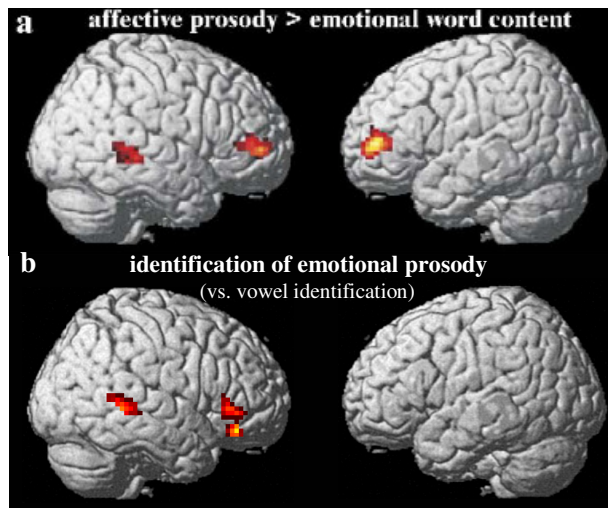


Figure 4: Task-dependent effects. (a) Increasing cerebral responses during explicit evaluation of emotional prosody at word level as compared to rating of emotional word content rendered on the cortex of a standard brain. Stimulus-driven activation during the same experiment is shown in Fig. 2. (b) Task-dependent effects at sentence level. Significant activations during categorization of emotional prosody as compared to the phonetic control task. Evaluation of emotional prosody yielded activation within the right post-STC and inferior frontal cortex.

During another experiment semantically neutral sentences (i.e. „The visitor reserved a room for Thursday“), spoken by actors in different emotional tones (happy, fearful, angry, sad, disgusted) served as stimulus material (Wildgruber et al., 2005). Participants either had to name the emotional category

expressed by prosody or they had to perform a phonetic control task (identification of the vowel following the first “a” in the sentence). In accordance with the findings during processing of single words, explicit evaluation of emotional prosody at sentence level was associated with activation of the right post-STC and the inferior frontal cortex (Fig. 4b).

A different pattern of task-dependent effects was observed, when subjects were asked to voluntarily modulate spatial attention while performing a gender discrimination task during a dichotic listening experiment. Anger prosody presented at the to-be-attended side yielded increasing activation within the medial orbitofrontal cortex (OFC) and the medial occipital cortex as compared to presentation of identical stimuli at the to-be-ignored ear (Sander et al., 2005, Fig. 5). The right amygdala and the bilateral mid-STC, in contrast, responded to anger prosody irrespective of the direction of spatial attention.

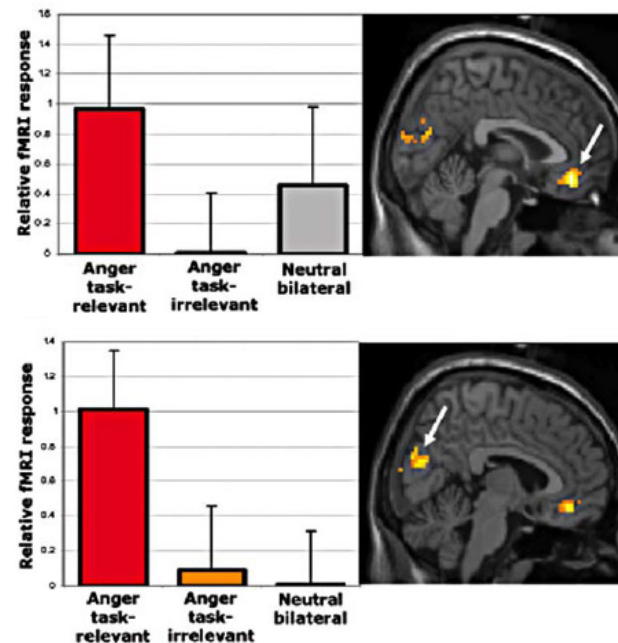


Figure 5: Spatial attention dependent effects. Meaningless utterances, pronounced with either angry or neutral prosody were presented in a dichotic listening paradigm. Participants were asked to selectively attend to either the left or right ear and performed a gender-decision on the voice heard on the target side. Increasing activation to the angry voice presented on the to-be-attended as compared to the to-be-ignored ear was observed within the medial orbitofrontal cortex (a) and the medial occipital cortex (b) (adapted from Sander et al., 2005).

Besides emotional information, speech prosody can also convey information about linguistic meaning (e.g. determining if a sentence is a statement, a question or a command). Experimental findings indicate that the contribution of the right and the left cerebral hemisphere to extraction of acoustic signals depends upon specific stimulus properties (Wildgruber et al., 2006; Reiterer et al., 2005, 2008). According to the acoustic-lateralization hypothesis slow changes of acoustic signals (e.g. modulations of prosody) are processed within the right hemisphere, whereas the left hemisphere is better suited to process rapid changes of acoustic signals (e.g. differentiation of speech sounds at the

level of syllables or phonemes). To further disentangle the impact of functional aspects (emotional vs. linguistic) from basic auditory processing, evaluation of emotional and linguistic prosody was compared using acoustically highly controlled stimuli. To this end, the intonation-contour of a semantically neutral sentence (“the scarf is in the chest”) was systematically manipulated by digital resynthesis. Participants were asked to differentiate pairs of these sentences either with respect to their emotional arousal (“which of the two sentences sounds more excited?”) or their sentence focus (“which of the two sentences is better suited to answer the question: where is the scarf?”). As compared to rest both tasks yielded bilateral frontotemporal activation including rightward lateralization at the level of the post-STC (Wildgruber et al., 2004). Analysis of task-effects, however, revealed activation of bilateral orbitofrontal cortices linked to emotional evaluation (Fig. 6).

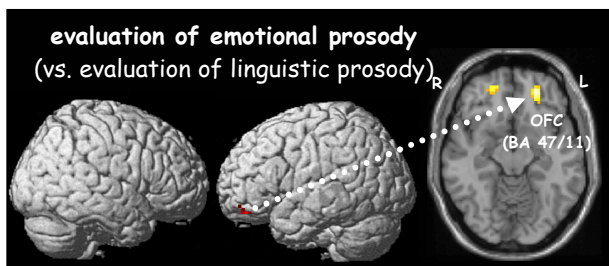


Figure 6: Task-dependent activation during evaluation of emotional prosody as compared to evaluation of linguistic prosody. Significant responses were rendered on the cortex of a standard brain (left) and a transversal slice (right) at the level of the highest activated voxel.

### 3. Model of emotional prosody processing

An analysis of effective connectivity revealed that the right post-STC is the most likely input region into the network of areas characterised by task-dependent activation (Ethofer et al., 2006b). This finding is in line with the assumption that this area subserves representation of meaningful prosodic sequences and receives direct input from primary and secondary acoustic regions. Moreover, the connectivity analysis indicated a flow of information along parallel projections from the right post-STC to the bilateral inferior frontal cortex. Taken together these findings indicate multiple successive processing stages, during recognition of emotional prosody following representation within the primary auditory cortex (A1). The first step, extraction of supra-segmental acoustic information, is associated with activation of predominantly right hemispheric primary and secondary acoustic regions. The second step, representation of meaningful supra-segmental acoustic sequences, is linked to posterior aspects of the right superior temporal sulcus. The third step, emotional judgment, is linked to the bilateral inferior-frontal cortex (IFC). Within this network, the projection from primary acoustic regions (A1) to the secondary acoustic representation within the mid-STC seems to be predominantly stimulus-driven (bottom-up effect), whereas further projections to the post-STC and the IFC depend upon focussing of attention towards explicit emotional evaluation (top-down effects). Considering implicit processing of emotional prosody, task-independent activation of the

amygdala has been observed (Sander et al., 2005; Ethofer et al., 2008). Additionally, top-down modulation depending on spatial attention has been observed within the medial OFC. Therefore, this region might contribute to processing of emotionally relevant information within the scope of attention even if no explicit emotional judgement is required. Moreover, the observed activation pattern within the OFC seems to indicate that this area might contribute to the inhibition of potentially distracting information from unattended sources. Presumably, the medial OFC receives its input from subcortical limbic regions such as the amygdala (Fig. 7).

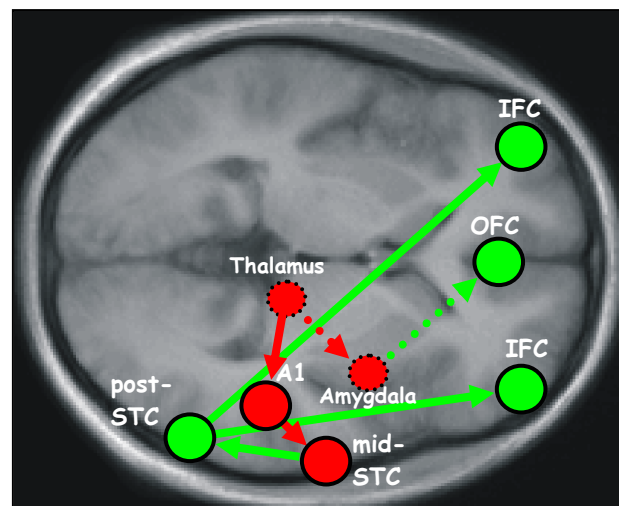


Figure 7: Model of emotional prosody processing. Processing steps during explicit evaluation (following signal transmission via thalamus and A1):

- 1) Secondary auditory representation (mid-STC)
- 2) Identification of prosodic sequences (post-STC)
- 3) Explicit emotional evaluation (IFC / OFC)

Flow of information during implicit processing

- 1) Thalamus
- 2) Amygdala
- 3) Medial orbitofrontal cortex (OFC)

Bottom-up modulation (stimulus-driven) is indicated with red arrows. Top-down modulation (task-dependent) is indicated with green arrows. A1 = primary auditory cortex, mid-STC = middle section of the superior temporal cortex, post-STC = posterior section of the STC, IFC = inferior frontal cortex, OFC = orbitofrontal cortex.

### 4. Cross modal integration of emotional signals

Considering bimodal integration, evaluation of audiovisual emotional signals yielded increased activation within the bilateral post-STC and the right thalamus as compared to either unimodal stimulation (Ethofer et al., 2006d, Kreifelts et al., 2007). Moreover, enhanced connectivity between the bilateral post-STC and voice-sensitive (mid-STC) as well as face-sensitive (fusiform face area) regions during processing of multimodal signals has been observed, which possibly depicts the mechanism of bimodal binding (Kreifelts et al., 2007). Presumably, the formation of such multimodal associations within the post-STC might also contribute to correct understanding of unimodal communicative signals.



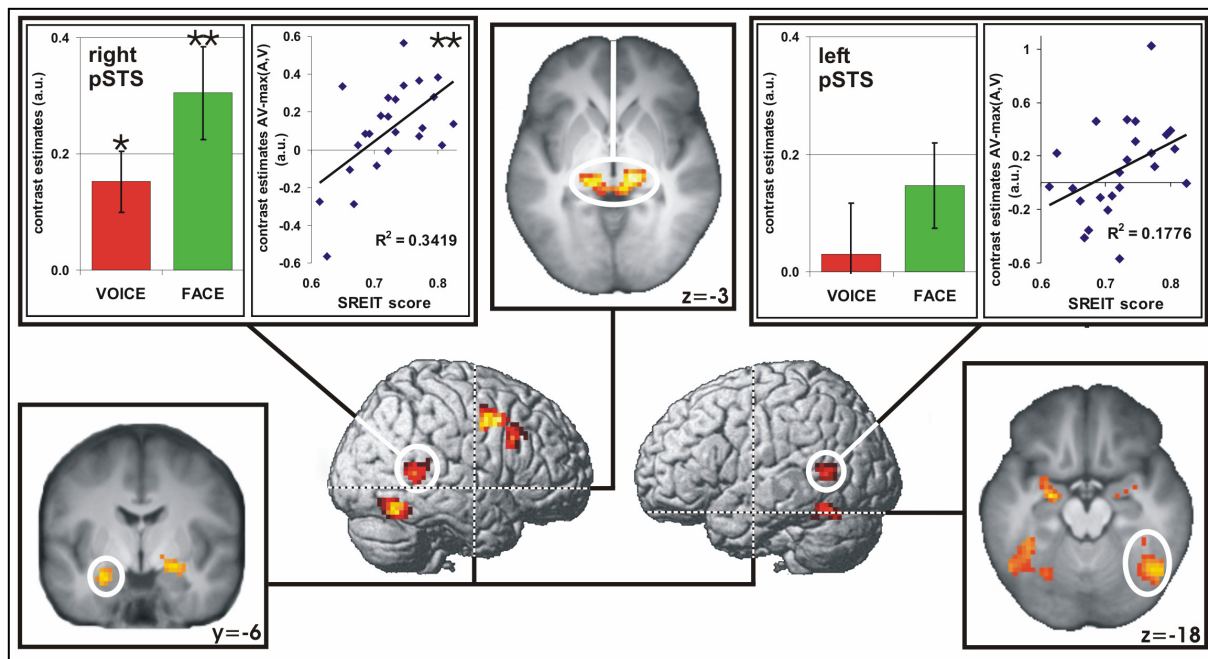


Figure 8: Activation during implicit audiovisual integration as compared to either unimodal stimulation. White circles mark regions of interest within bilateral post-STC, right thalamus, right fusiform gyrus and left amygdala. Vertical-bar diagrams depict voice sensitivity (red) and face sensitivity (green) within the respective regions. Only the right post-STC showed a correlation of individual response patterns and a measure of trait emotional intelligence (SREIT scores) as well as voice and face sensitivity.

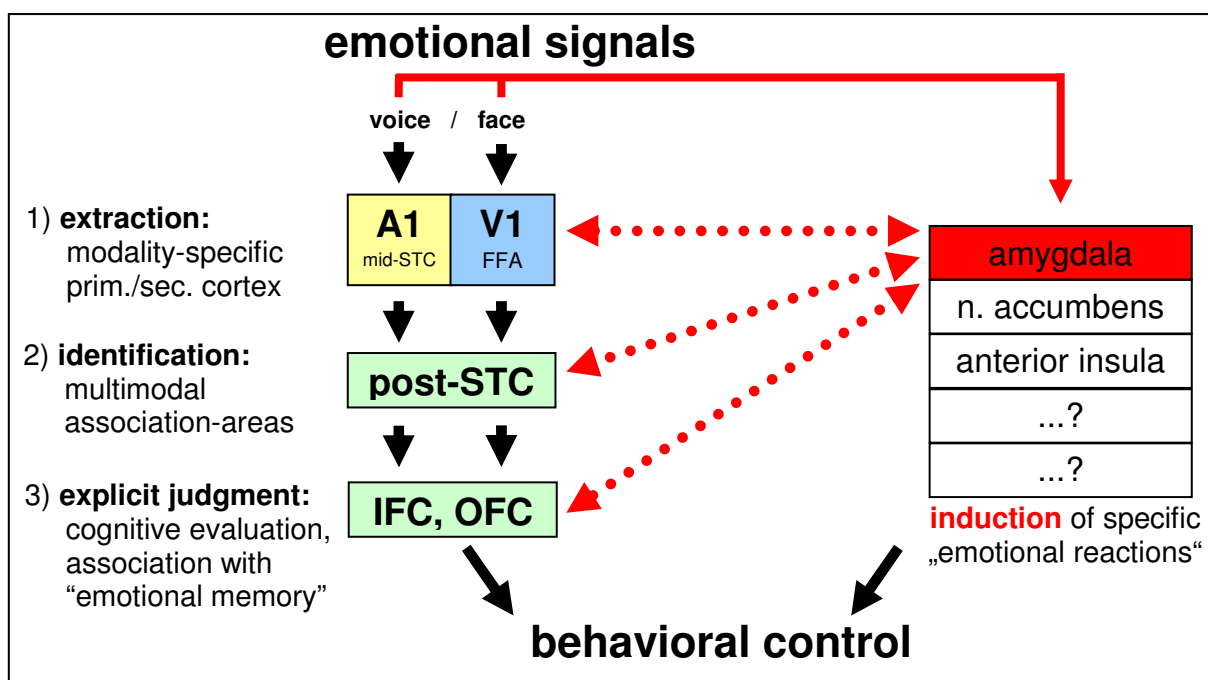


Figure 9: Cross-modal integration of emotional communicative signals. (1) Extraction of different communicative signals is performed within the respective modality-specific primary and secondary cortices (mid-STC = middle section of the superior temporal cortex, FFA = fusiform face area). (2) Identification of meaningful sequences relies on multimodal associations areas (post-STC = posterior section of the STC). (3) As a third step, explicit emotional judgements are subserved within the inferior frontal cortex (IFC) and the orbitofrontal cortex (OFC). Moreover, emotional signals can yield an automatic induction of emotional reactions that is linked to specific subcortical regions. Besides a direct connection from the thalamus to the respective subcortical areas (red line), presumably, both neural pathways are interconnected at various levels (red dots) allowing for a complex reciprocal interaction of cognitive evaluation (explicit) and automatic processing (implicit) of emotional signals.

In these instances missing sensory information might be complemented on the basis of established associations from memory to determine the “meaning” of the signal. During social interaction, however, the emotional connotations of communicative signals are usually not explicitly judged. Rather, highly automatic monitoring of emotional information conveyed by various channels of communication is permanently required. A variety of experimental data indicate different cerebral pathways to be involved in explicit and implicit processing of emotional signals. Limbic brain structures sensitive to emotional information (e.g. amygdala, nucleus accumbens), exhibit a more pronounced response during implicit stimulus processing as compared to explicit and cognitively controlled evaluation (Hariri et al., 2003; Lange et al., 2003). Based on functional imaging studies it has been postulated, that dorsolateral and lateral orbitofrontal areas are involved in inhibition of limbic activation during explicit emotional judgments (Blair et al., 2007; Mitchell et al., 2007). Increased activity within the medial frontal cortex, in contrast, has been associated with enhanced affective evaluation contributing to selection of emotional significant information in accordance with individual motivational goals and current behavioral demands (Kringelbach, 2005; Phillips et al., 2003). Taken together these findings suggest a predominant role of subcortical limbic regions for implicit emotional processing and a stronger involvement of cortical regions (various frontal areas as well as modality-specific sensory areas) during explicit and cognitively controlled processing of emotional stimuli.

Considering cross modal integration, facial expressions were rated as being more fearful when presented concomitant with fearful prosody. These changes due to implicitly processed fearful prosody were correlated with enhanced activation of the left amygdala (Ethofer et al., 2006c). In a recent experiment, implicit cross-modal integration was evaluated while subject had to perform a gender discrimination task (Kreifelts et al., 2008). Bimodal stimulation yielded increasing activation of post-STC, thalamus, amygdalae and fusiform gyri (Fig.8). Among these areas, however, solely the right post-STC displayed a positive correlation of the individual hemodynamic integration effect and a measure of trait emotional intelligence as well as voice and face sensitivity. Cumulating evidence, thus, indicates that the post-STC might serve as an essential interface between perceptual integration and social cognition (Fig. 9).

## 5. References

- [1] Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature* 403, 309-312.
- [2] Blair, K.S., Smith B.W., Mitchell, D.G.V., Morton, J., Vythilingam, M., Pessoa, L., Fridberg, D., Zametkin, A., Nelson, E.E., Drevets, W.C., Pine D.S., Martin, A., Blair, R.J.R. (2007) Modulation of emotion by cognition and cognition by emotion. *NeuroImage* 35, 430-440.
- [3] Ethofer, T., Erb, M., Anders, S., Wiethoff, S., Herbert, C., Saur, R., Grodd, W., Wildgruber, D. (2006a) Effects of prosodic emotional intensity on activation of associative auditory cortex, *NeuroReport* 17, 249-253.
- [4] Ethofer, T., Anders, S., Erb, M., Herbert, C., Wiethoff, S., Kissler, J., Grodd, W., Wildgruber, D. (2006b) Cerebral pathways in processing of emotional prosody: A dynamic causal modelling study. *NeuroImage* 30, 580-587.
- [5] Ethofer, T., Pourtois, G., Wildgruber, D. (2006c) Investigating audiovisual integration of emotional information in the human brain. *Progress in Brain Research* 156, 345-361.
- [6] Ethofer, T., Anders, S., Erb, M., Droll, C., Royen, L., Saur, R., Reiterer, S., Grodd, W., Wildgruber, D. (2006d) Impact of voice on emotional judgement of faces: an event-related fMRI study. *Human Brain Mapping* 27, 707-714.
- [7] Ethofer, T., Wiethoff, S., Anders, S., Kreifelts, B., Grodd, W., Wildgruber, D. (2007) The voices of seduction: Cross-gender effects in processing of erotic prosody. *Social Cognition and Affective Neuroscience* 2, 334-337.
- [8] Ethofer, T., Kreifelts, B., Wiethoff, S., Wolf, J., Grodd, W., Vuilleumier, P., Wildgruber, D. (2008) Repetition suppression in orbitofrontal cortex is modulated by angry prosody, *submitted*.
- [9] 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.
- [10] Hariri, A.R., Mattay, V.S., Tessitore, A., Fera, F., Weinberger, D.R., 2003. Neocortical modulation of the amygdala response to fearful stimuli. *Biol Psychiatry* 53, 494-501.
- [11] Kreifelts, B., Ethofer, T., Grodd, W., Erb, M., Wildgruber, D. (2007) Audiovisual integration of emotional signals in voice and face: an event-related fMRI study. *NeuroImage* 37, 1445-1456.
- [12] Kreifelts, B., Ethofer, T., Huberle, E., Grodd, W., Wildgruber, D. (2008) Association of trait emotional intelligence and individual fMRI-activation patterns during emotional perception, *submitted*.
- [13] Kringelbach, M.L. (2005) The human orbito-frontal cortex: linking reward to hedonic experience. *Nature Review Neuroscience* 6, 691-702.
- [14] Lange, K., Williams, L.M., Young, A.W., Bullmore, E.T., Brammer, M.J., Williams, S.C., Gray, J.A., Phillips, M.L. (2003). Task instructions modulate neural responses to fearful facial expressions. *Biol Psychiatry* 53, 226-232.
- [15] Mitchell, D.G.V., Nakic, M., Fridberg, D., Kamel, N., Pine, D.S., Blair, R.J.R. (2007) The impact of processing load on emotion. *NeuroImage* 34, 1299-1309.
- [16] Phillips, M.L., Drevets, W.C., Rauch, S.L., Lane, R. (2003). Neurobiology of Emotion Perception I: The Neural Basis of Normal Emotion Perception. *Biological Psychiatry* 54, 504-514.
- [17] Sander, D., Grandjean, D., Pourtois, G., Schwartz, S., Seghier, M., Scherer, K.R., Vuilleumier, P. (2005). Emotion and Attention Interactions in Social Cognition: Brain Regions Involved in Processing Anger Prosody. *NeuroImage*. 28, 848-858.
- [18] Schirmer, A., Kotz, S.A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences* 10, 24-30.

- [19] Wildgruber, D., Hertrich, I., Riecker, A., Erb, M., Anders, S., Grodd, W., Ackermann, H. (2004) Distinct Frontal Regions Subserve Evaluation of Linguistic and Affective Aspects of Intonation. *Cerebral Cortex* 14, 1384-1389.
- [20] 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.
- [21] Wildgruber, D., Ackermann, H., Kreifelts, B., Ethofer, T. (2006) Processing of Linguistic and Emotional Prosody: fMRI Studies. *Progress in Brain Research* 156, 249-268.
- [22] Wiethoff, S., Wildgruber, D., Becker, H., Anders, S., Herbert, C., Grodd, W., Ethofer, T. (2008) Cerebral processing of emotional prosody - Influence of acoustic parameters and arousal. *NeuroImage* 39, 885-893.