

Cerebral Processing of Linguistic and Emotional Prosody: fMRI Studies

D. Wildgruber^{1,2}, H. Ackermann³, B. Kreifelts³, T. Ethofer^{1,2}

¹*Department of Psychiatry*

²*Section MR of CNS, Department of Neuroradiology*

³*Department of General Neurology, Hertie Institute for Clinical Brain Research
University of Tübingen, Germany*

Manuscript submitted to *Progress in Brain Research*, January 31, 2006

Corresponding author: Dirk Wildgruber, M.A., M.D.
Department of Psychiatry
University of Tübingen
Osianderstr. 24
72076 Tübingen, Germany
Phone: +49-7071-2987726
Fax: +49-7071-294371
e-mail: dirk.wildgruber@med.uni-tuebingen.de

No. of color figures: 4 (Fig. 1, Fig. 2, Fig. 3, Fig. 4)

Key Words:

Affect, Communication, Emotion, fMRI, Intonation, Language, Lateralization, Prosody

Running Title:

Cerebral Processing of linguistic and emotional prosody

Abstract:

During acoustic communication in humans, information about a speaker's emotional state is predominantly conveyed by modulation of the tone of voice (emotional prosody). Based on lesion data, a right-hemisphere superiority for cerebral processing of emotional prosody has been assumed. However, the available clinical studies do not yet provide a coherent picture with respect to inter-hemispheric lateralization effects of prosody recognition and intra-hemispheric localization of the respective brain regions. In order to further delineate the cerebral network engaged in the perception of emotional tone, a series of experiments was carried out based upon functional magnetic resonance imaging (fMRI). The findings obtained from these investigations allow for the separation of three successive processing stages during recognition of emotional prosody: (1) extraction of supra-segmental acoustic information predominantly subserved by right-sided primary and higher-order acoustic regions; (2) representation of meaningful supra-segmental acoustic sequences within posterior aspects of the right superior temporal sulcus; (3) explicit evaluation of emotional prosody at the level of the bilateral inferior-frontal cortex. Moreover, implicit processing of emotional prosody seems to be bound to subcortical regions mediating automatic induction of specific emotional reactions such as activation of the amygdala in response to fearful stimuli.

As concerns lower-level processing of the underlying supra-segmental acoustic cues, linguistic and emotional prosody seem to share the same right-hemisphere neural resources. Explicit judgment of linguistic aspects of speech prosody, however, appears to be linked to left-sided language areas whereas bilateral orbito-frontal cortex has been found involved in explicit evaluation of

emotional prosody. These differences in hemispheric lateralization effects might explain that specific impairments in non-verbal emotional communication subsequent to focal brain lesions are relatively rare clinical observations as compared to the more frequent aphasic disorders.

Introduction:

During social interactions among humans, transfer of information does not only depend upon the words we use. Rather, in numerous situations it seems to be much more important *how* we utter them (Mehrabian, 1972). Emotional states, attitudes (i.g., sympathy, dominance, politeness), and intentions often are predominantly expressed by the modulation of the tone of voice (emotional prosody). For example, if your head of department comes around and says with an angry intonation "I have just been reading your report. We have to talk about it right now" you will certainly get a fairly different impression of his intentions as if he would produce the same sentences in a friendly and happy manner. As concerns the cerebral correlates of prosody processing, observations in patients suffering from focal brain lesions indicate that the well-established left-hemisphere dominance for language comprehension does not extend to the perception of emotional tone (Hughlings-Jackson, 1879; Pell and Baum, 1997; Schmitt et al., 1997; Baum and Pell, 1999; Borod et al., 2001, 2002; Adolphs, 2002; Charbonneau et al., 2003; Wildgruber and Ackermann 2003; Ackermann et al., 2004). According to an early neuroanatomical model proposed by Ross (1981), prosodic information is encoded within distinct right-sided perisylvian regions that are organized in complete analogy to the left-sided language areas. Expression of emotional prosody, thus, is believed

to depend upon the Broca-homologue within the right inferior frontal cortex whereas comprehension of intonational information is presumed to be bound to the right superior temporal region (Wernicke-homologue). However, the empirical evidence for this model was based on a few case reports only, and more systematic investigations yielded rather discrepant results. The majority of lesion studies seem to be compatible with the assumption that the right-hemisphere posterior perisylvian cortex is highly important for the comprehension of speech melody (Heilman et al., 1975, 1984; Darby, 1993; Starkstein et al., 1994; Adolphs et al., 2001; Borod et al. 2002). However, various clinical examinations indicate a widespread network of - partially bilateral - cerebral regions including the frontal cortex (Hornack et al., 1996, 2003; Breitenstein et al., 1998; Rolls, 1999; Adolphs et al., 2002) and the basal ganglia (Cancellier and Kertesz 1990; Weddel, 1994; Peper and Irle, 1997; Breitenstein et al., 1998, 2001; Pell and Leonard, 2003) to contribute to the processing of emotional intonation. In line with these findings, several neuroimaging studies reported rightward lateralization of hemodynamic activation within temporal regions (Buchanan et al., 2000; Wildgruber et al., 2002, 2005; Mitchell et al., 2003; Kotz et al., 2003; Grandjean et al., 2005) and revealed additional – partially bilateral – responses within the frontal cortex (George et al., 1996; Imaizumi et al., 1997; Buchanan et al., 2000, Wildgruber et al., 2002, 2004, 2005; Kotz et al., 2003), the anterior insula (Imaizumi et al., 1997; Wildgruber et al., 2002, 2004) and the basal ganglia (Kotz et al., 2003) during recognition of emotional intonation. The considerable differences in lateralization and localization of the relevant lesion sites as well as hemodynamic activation spots, however, do not yet allow for an indisputable determination of the neural substrates of prosody processing. Presumably, the discrepancies of the available data are due to differences in the methods used such as stimulus selection,

task and control conditions. In order to further clarify to which extent specific neural structures subserve different facets of the comprehension of emotional prosody, our research group conducted a variety of experiments based on functional magnetic resonance imaging (fMRI), a technique that can be used for the non-invasive evaluation of task-related hemodynamic cerebral responses at a highly spatial (ca. 0.5 mm, Menon and Goodyear, 1999) and moderately temporal (< 1 second, Wildgruber et al., 1997) resolution. Specifically, these studies were designed to delineate the neural substrates underlying distinct facets of prosody processing: (a) extraction of supra-segmental acoustic information, (b) representation of meaningful prosodic sequences, (c) explicit judgment of emotional as compared to linguistic information, (d) connectivity between the neural structures involved, and (e) implicit processing of emotional prosody.

Extraction of supra-segmental acoustic Information

At the perceptual level, emotional tone is characterized by the modulation of loudness (acoustic correlate: sound intensity), pitch (fundamental frequency variation), speech rhythm (duration of syllables and pauses), and voice quality or timbre (distribution of spectral energy) across utterances (Lehiste, 1970; Ackermann et al., 1993; Murray and Arnott, 1993; Banse and Scherer, 1996; Cutler et al., 1997; Bachorowski and Owren, 2003; Sidtis and Van-Lancker-Sidtis, 2003; Scherer et al., 2003). These supra-segmental features are imposed upon the sequence of speech sounds (segmental structure) of verbal utterances. According to the acoustic lateralization hypothesis (Fig. 1a), the encoding of supra-segmental parameters of the speech signal (rather slow shifts > 100 ms) is predominantly bound to right-

hemisphere structures whereas rapid transitions (< 50 ms), contributing to the differentiation of the various speech sounds at segmental level (i.e., phonemes, syllables), are mainly processed within contralateral areas (Ivry and Robertson, 1998; Belin et al., 1998; Zatorre et al., 2001, 2002; Meyer et al., 2002; Poeppel et al., 2004). These acoustic laterality effects have been supposed to explain the differential hemispheric dominance patterns of language (left Hemisphere) and music processing (right hemisphere) (Wildgruber et al., 1996, 1998, 2001, 2003; Belin 1998; Ivry and Robertson 1998; Zatorre et al., 2002; Hugdahl and Davidson, 2003; Poeppel, 2004; Ackermann et al., 2006). In order to further separate the neural structures subserving the extraction of basic acoustic properties of speech prosody from those which respond to the conveyed emotional “meaning”, a series of fMRI experiments was conducted. More specifically, the following hypotheses were explored:

- a) Lateralization of hemodynamic responses during passive listening to trains of noise bursts depends upon stimulus frequency.
- b) Extraction of specific acoustic parameters (signal duration, fundamental frequency) is associated with different activation patterns at the level of primary and higher-order acoustic regions.
- c) Expressiveness of emotional prosody enhances the hemodynamic responses of voice-sensitive areas within the right as compared to corresponding regions within the left hemisphere.

The first experiment encompassed a simple passive listening condition. Trains of noise bursts (clicks) were presented at different rates (2.0, 2.5, 3.0, 4.0, 5.0, 6.0 Hz) to eight healthy right-handed subjects (4 males and 4 females, age 19-32 years)

during fMRI measurements. The clicks had been produced originally by striking a pen against a table. Each acoustic sequence of a given click rate had a duration of 6 seconds. Altogether, 90 trains (6 rates x 15 repetitions) were presented in pseudo-randomized order. During passive listening to these stimuli, significant hemodynamic responses across all different presentation rates emerged within the superior temporal gyrus of both sides, right-hemisphere putamen and the tectum. Moreover, parametric analysis revealed lateralized rate-dependent responses within the anterior insular cortex. During presentation of the click trains at slow rates, the right anterior insula showed the highest activation levels. Furthermore, the hemodynamic responses of this region displayed a decline of amplitude in parallel with an increase of stimulation frequency. By contrast, an opposite relationship emerged within the left anterior insular cortex (Ackermann et al., 2001). This double dissociation of rate-response functions between the two hemispheres is in a very good accordance with the acoustic lateralization hypothesis (Fig. 1b).

Seventeen healthy volunteers (8 males, 9 females, age 18-31 years) participated in a second experiment that investigated discrimination of duration and pitch values at different levels of difficulty. Complex sounds consisting of four formant frequencies (500, 1500, 2500, 3500 Hz), manipulated either in duration (100-400 ms) or in fundamental frequency (100-200 Hz, realized by rhythmic intensity fluctuations throughout the signal), served as stimuli. Sequences of two signals were presented to both ears each, and subjects either had to detect the longer duration (duration task) or the higher pitch (pitch task), respectively. The behavioural data showed comparable hit scores (mean values about 75%) with increasing accuracy rates in correlation to rising physical difference between the two acoustic signals for both, pitch and duration discrimination (Fig 1c). As compared to baseline at rest, both tasks

yielded bilateral activation of frontal, temporal and parietal regions including primary and secondary acoustic cortices as well as the working memory network (Fig. 1c). A lateralization analysis, i.e., comparison of each hemisphere with the contralateral side on a voxel-by-voxel basis, revealed, however, lateralization effects toward the left side within insular and temporal cortex during both tasks. Even more noteworthy, a parametric analysis of hemodynamic responses showed an increase of activation within the right temporal cortex in parallel with the differences in sound properties of the stimulus pairs. This positive linear relationship emerged both during the duration and the pitch task. Moreover, a comparison with the contralateral hemisphere revealed significant lateralization effects of the parametric responses towards the right superior temporal sulcus during discrimination of stimulus duration (Reiterer et al., 2005). Slowly changing and highly different acoustic stimuli, thus, seem to be predominantly processed within the right hemisphere, whereas detection of rapid changes or rather slight signal differences might be linked to the left hemisphere.

The findings of these first two experiments indicate that differences in basic acoustic properties have a strong impact on brain activation patterns. In a third study, twelve healthy right-handed subjects (7 males, 5 females, aged 19-29 years) were asked to judge in two separate sessions the emotional valence of either word content or prosody of altogether 162 German adjectives spoken in a happy, angry or neutral tone. Intonations of these different emotional categories differ in various acoustic properties (Banse and Scherer, 1996). To disambiguate more specific effects of emotional expressiveness from extraction of low-level acoustic parameters, mean and variation of volume intensity and fundamental frequency were included in the statistical models as nuisance variables. During both tasks, a linear correlation between hemodynamic responses and prosodic emotional expressiveness emerged

within the middle part of the bilateral superior temporal sulcus (mid-STS). Responses of right-hemisphere mid-STS showed higher amplitudes, larger extension and a stronger dependency on emotional intensity than those of the contralateral side (Fig. 1d). Similar response patterns were found both for explicit and implicit processing of emotional prosody (Ethofer et al., 2006). These observations support the assumption that the mid-STS region contributes to the encoding of emotionally salient acoustic stimuli independent from task-related attentional modulation (Grandjean et al., 2005).

In summary, these findings, related to the acoustic level of prosody processing, indicate extraction of supra-segmental acoustic information to be predominantly subserved within right-sided primary and higher-order acoustic brain regions including mid-STS and anterior insula.

[Please place Fig. 1 near here]

Representation of meaningful prosodic sequences

According to the neuroanatomical model proposed by Elliot Ross, the Wernicke-homologue region bound to the posterior aspects of right-hemisphere superior temporal gyrus represents the key area for the comprehension of prosodic sequences (Ross, 1981). An important role of the right posterior perisylvian cortex for comprehension of speech melody has been confirmed in various clinical examinations (Heilman et al., 1975, 1984; Darby, 1993; Starkstein et al., 1994; Borod et al. 2002). In some studies on the comprehension of emotional information, however, the valence of emotional expression has been reported to influence

lateralization of cerebral responses (Canli et al., 1998; Davidson et al., 1999; Murphy et al., 2003). According to the valence hypothesis, rightward lateralization of prosody processing only holds true for negative emotions, whereas comprehension of happy stimuli is ascribed to the left hemisphere (Fig. 2). As concerns speech intonation, several clinical examinations failed to show any interactions between hemispheric lateralization and emotional valence (Pell, 1998; Baum and Pell 1999; Borod et al., 2002; Kucharska-Pietura et al., 2003). Considering functional imaging data, however, distinct cerebral activation patterns bound to specific emotional categories such as disgust, anger, fear or sadness have been observed during perception of facial emotional expressions (Sprengelmeyer et al., 1998; Kesler-West et al., 2001; Phan et al., 2002; Murphy et al., 2003). Several studies have corroborated the notion that responses of the amygdalae are specifically related to facial expressions of fear (Morris et al., 1996, 1998; Adolphs 2002; Phan et al., 2002) whereas facial expressions of disgust seem to elicit activation of the anterior insula (Sprengelmeyer et al., 1998; Phillips et al., 1998; Calder et al., 2001; Phan et al. 2002; Wicker et al., 2003). Fear-specific responses of the amygdalae have also been reported in association with vocal emotional expressions (Phillips et al., 1998; Morris et al., 1999) whereas the predicted disgust-related activation of the anterior insula has not been observed in a prior PET experiment (Phillips et al. 1998). It is unsettled, thus, to which extent lateralization and exact localization of cerebral activation during comprehension of emotional prosody is linked to specific emotional categories.

Based on the aforementioned clinical and neuroimaging studies, presumably, there are cerebral regions – including the right posterior temporal cortex - that contribute to comprehension of emotional prosody independent of any specific emotional content, whereas other regions – including the amygdala and anterior

insula - are selectively linked to comprehension of specific emotional categories. In order to separate these components, 100 short German declarative sentences with emotionally neutral content (such as “Der Gast hat sich für Donnerstag ein Zimmer reserviert [*The visitor reserved a room for Thursday*], “Die Anrufe werden automatisch beantwortet [*Phonecalls are answered automatically*]”) were randomly ascribed to one of five different target emotions (happiness, anger, fear, sadness or disgust). A professional actress and an actor produced these test materials expressing the respective emotion by modulation of affective intonation. Verbal utterances were presented to 10 healthy subjects (5 males, 5 females, age: 21-33 years) under two different task conditions during fMRI. As an identification task, subjects were asked to name the emotion expressed by the tone of voice whereas the control condition (phonetic task) required the detection of the vowel following the first /a/ in each sentence. Similarly to the emotion recognition task, vowel identification also included a forced choice selection from five alternatives, i.e., the vowels (/a/, /e/, /i/, /o/, /u/). Under both conditions, participants were asked to give a verbal response as quickly as possible and they were provided with a list of possible response alternatives prior to testing. Since both tasks require evaluation of completely identical acoustic stimuli and involve very similar response mechanisms, comparison of the respective hemodynamic activation patterns should allow for the separation of task-specific cerebral responses independently of stimulus characteristics and unspecific task components. In order to delineate cerebral structures contributing to the recognition of emotional prosody independent of specific emotional categories, responses during the identification of emotional prosody across all emotional categories were compared to the phonetic control condition. To disentangle patterns of cerebral activation related to comprehension of specific

emotional categories, each emotional category was compared against the others.

The main goal of the study, thus, was to evaluate the following two hypotheses:

- (a) A network of right-hemisphere areas including the posterior temporal cortex supports identification of affective intonation independent of specific emotional information conveyed.

- (b) Perception of different emotional categories is associated with specific brain regions, i.e., response localization varies with emotion type. Specifically, fear-specific responses are linked to the amygdalae and disgust-specific responses to the anterior insula.

During the fMRI experiment, subjects correctly identified the emotional tone at a slightly lower rate (mean: 75.2 ± 7.9 %) as compared to the vowel detection task (mean: 83.4 ± 7.0 %, $p < 0.05$). The accuracy scores for happy (90%), angry (82%) and sad (84%) expressions reached comparable levels whereas fearful (51%) and disgusted (57%) expressions were identified at significantly lower rates ($p < 0.05$). These differences in performance are in good accordance with prior observations and might be related to differences in recognizability of the acoustic cues of the various emotions (Banse and Scherer, 1996). Response times for the emotional task (mean: 4.3 ± 0.9 seconds) showed no significant differences as compared to the phonetic task (mean: 4.1 ± 1.0 seconds) indicating comparable levels of task difficulty. Cerebral responses obtained during both tasks, as compared to the rest condition, yielded a bilateral network of hemodynamic activation at the level of cortical and subcortical regions including frontal, temporal and parietal cortex, thalamus and cerebellum. To identify brain regions specifically contributing to the

encoding of emotional intonation, the respective activation patterns were directly compared to the responses obtained during phonetic processing of the identical acoustic stimuli (Wildgruber et al., 2005). Using this approach, responses within two activation-clusters, localized within the right posterior superior temporal sulcus (BA 22/42) and the right inferior frontal cortex (BA 45/47), could be assigned to recognition of emotional prosody (Fig. 2b). No significant impact of emotional valence or specific emotional categories on the distribution of brain activation could be observed. Therefore, the results of the current study do not support, in line with prior functional imaging (Buchanan et al., 2000; Wildgruber et al., 2002; Mitchell et al., 2003; Kotz et al., 2003) and recent lesion studies (Pell, 1998; Baum and Pell 1999; Borod et al., 2002; Kucharska-Pietura et al., 2003), the hypothesis of valence-specific lateralization effects during processing of emotional intonation.

The observed hemodynamic responses, however, indicate a task-dependent and stimulus-independent contribution of the right posterior STS (BA 22/42) and the right inferior frontal cortex (BA 45/47) to processing of supra-segmental acoustic information irrespective of specific emotional categories. We assume, therefore, that the representation of meaningful supra-segmental acoustic sequences within these areas should be considered a second step of prosody processing. A further experiment was designed in order to evaluate the contribution of posterior STS and inferior frontal cortex to the processing of emotional prosody as compared to evaluation of linguistic prosody.

[Please place Fig. 2 near here]

Explicit Judgment of emotional prosody

As concerns its communicative functions, speech prosody serves a variety of different linguistic as well as emotional purposes (Ackermann et al., 1993, 2004; Baum and Pell, 1999). Among others, it is used to specify linguistic information at the word (**content** vs. **content**) and sentence level (question vs. statement intonation: "It is new?" vs. "It is new!"; location of sentence focus: "**he** wrote this letter " vs. "he wrote **this** letter"), and conveys information about a speaker's personality, attitude (i.e. dominance, submissiveness, politeness etc.) and emotional state (Fig. 3). Based on lesion studies, the functional lateralization hypothesis proposes linguistic prosody to be processed within the left hemisphere whereas emotional tone is bound to contralateral cerebral structures (Van Lancker, 1980; Heilman et al., 1984; Behrens, 1985; Emmorey, 1987; Pell and Baum 1997; Borod et al., 1998, 2002; Geigenberger and Ziegler, 2001; Schirmer et al., 2001; Charbonneau et al., 2003). In order to disentangle the functional and the acoustic level of prosody processing, sentences varying in linguistic accentuation (sentence focus) as well as emotional expressiveness were generated by systematic manipulations of the fundamental frequency contour of the simple declarative German sentence "Der Schal ist in der Truhe" ("The scarf is in the chest"). With its focus on the second word, this utterance represents an answer to the question "What is in the chest?". Shifting the accent to the final word, the sentence provides information about where the scarf is. This prosodic distinction is realized by distinct pitch patterns characterized by F0-peaks on the accented syllables (Cutler et al., 1997). As a first step, a series of five F0 contours was generated extending from a clear-cut focus on the second to an accent on the final word (Fig. 3b). On the basis of each of these five focus patterns, second,

five additional variations were generated differing in pitch range across the whole sentence. These global variations are perceived as modulations of emotional expressiveness. Sentences with broader F0 range clearly sound more excited (Banse and Scherer, 1996; Pihan et al. 1997). Ten healthy right-handed participants (6 males, 4 females, age: 20-35 years) were asked to perform two different discrimination tasks during pair-wise presentation of these acoustic stimuli. In two different sessions of the experiment they had to answer one of the following questions: (a) "Which of the two sentences is better suited as a response to the question: Where is the scarf?" (discrimination of linguistic prosody) and (b) "Which of the two sentences sounds more excited?" (discrimination of emotional expressiveness). Since both conditions require the evaluation of completely identical acoustic signals, the comparison of hemodynamic responses obtained during the two different runs allows for the separation of task-specific responses independent of stimulus characteristics. This experiment was primarily designed to explore the following two alternative hypotheses:

- a) Lateralization effects during prosody processing are strongly bound to acoustic properties of the relevant speech signal: Since comprehension of linguistic as well as emotional prosody relies upon the extraction of supra-segmental features, a rightward lateralization must be expected during both conditions (acoustic lateralization hypothesis)
- b) Linguistic prosody is processed within left-sided speech areas whereas comprehension of emotional prosody must be expected to be bound to the right hemisphere (functional lateralization hypothesis)

The obtained behavioral data clearly show that the participants were able to discriminate the patterns of linguistic accentuation and emotional expressiveness at similar levels of accuracy (linguistic discrimination: $82\% \pm 14\%$, emotional discrimination $78\% \pm 11\%$). Therefore, a comparable level of difficulty for both tasks can be assumed. As compared to the baseline at rest, both conditions yielded bilateral hemodynamic responses within supplementary motor area, anterior cingulate gyrus, superior temporal gyrus, frontal operculum, anterior insula, thalamus and cerebellum. Responses within the dorso-lateral frontal cortex (BA 9/45/46) showed lateralization effects toward the right side during both tasks (Wildgruber et al., 2004). In order to identify brain regions specifically contributing to the processing of linguistic or emotional intonation, the respective activation patterns were directly compared with each other.

During the linguistic task, significantly stronger activation was observed within the left inferior frontal gyrus (BA 44/45 = Broca's Area). By contrast, the affective condition yielded significant bilateral hemodynamic responses within orbitofrontal cortex (BA 11/47) as compared to the linguistic task (Fig. 3c). Comprehension of linguistic prosody requires analysis of the lexical, semantic and syntactic aspects of pitch modulation patterns. Activation of left inferior frontal cortex (Broca's area) concomitant with the discrimination of linguistic accents indicates that at least some of these operations might be housed within the anterior perisylvian language areas. In line with this assumption, native speakers of Thai, a tone language, showed activation of the left inferior frontal region during discrimination of linguistically relevant pitch patterns in Thai words. This activity was absent in English speaking subjects listening to identical stimuli (Gandour et al., 1998). Moreover, clinical observations support the assumption of a specific contribution of the left hemisphere

to the comprehension of linguistic aspects of intonation. For example, Heilman et al. (1984) found patients suffering from focal left-sided brain lesions to produce significantly more errors in a linguistic prosody identification task as compared to the recognition of affective intonation whereas damage to the right hemisphere was associated with a similar profile of deficits in both tasks. Furthermore, Emmorey (1987) observed impaired discrimination of stress contrasts between noun compounds and noun phrases after damage to the left hemisphere whereas patients with right-sided lesions performed as well as normal control subjects. Predominant disturbance of linguistic prosody comprehension concomitant with relatively preserved processing of emotional intonation in patients with damage to the left hemisphere has also been reported by Pell and Baum (1997) as well as Geigenberger and Ziegler (2001).

Discrimination of emotional expressiveness yielded a significant increase of hemodynamic responses within the bilateral orbito-frontal cortex (BA 11/47) as compared to the linguistic task indicating, thus, a specific contribution of this region to the evaluation of emotional aspects of verbal utterances conveyed by the tone of speech. On the basis of neuro-anatomical considerations, e.g. reciprocal fiber connections to sensory cortices and limbic regions, this region might serve as a substrate for the judgment of emotional stimuli independent of the stimulus modality (Price, 1999). Accordingly, activation of the orbito-basal frontal cortex has been observed in preceding functional imaging studies during perception of emotional intonation (George et al., 1996; Wildgruber et al., 2002), emotional facial expressions (Nakamura et al., 1999, Blair et al., 1999) and affective gustatory judgments (Small et al., 2001). Moreover, patients suffering from unilateral focal damage to this area displayed impaired identification of emotional face and voice expressions whereas

performance in non-emotional control tasks (i.e. discrimination of unfamiliar voices and recognition of environmental sounds) was found uncompromised (Hornak et al., 1996, 2003; Rolls 1999). These observations, in line with the results of the present study, support the assumption that orbito-frontal areas contribute to the explicit evaluation of emotional information conveyed by different communicational channels. Blair and Cipolatti supposed this region to be critically involved in building associations between the perceived emotional signals and an emotional episodic memory. In patients suffering from lesions of orbito-frontal cortex, pronounced abnormalities of social behavior have been observed (Levin et al., 1991; Blair and Cipolatti, 2000; Wildgruber et al., 2000), resulting, conceivably, from compromised associations between actual environmental stimuli with emotional memory traces.

In conclusion, hemispheric specialization for higher-level processing of intonation contours has been found to depend – at least partially - upon the functional role of the respective acoustic signals within the communication process: Comprehension of linguistic aspects of speech melody relies predominantly upon left-sided perisylvian language areas whereas the evaluation of emotional signals, independent of modality and emotion type, is bound to bilateral orbitofrontal regions. As a third step of prosody processing, thus, explicit evaluation of emotional prosody seems to be associated with bilateral inferior aspects of frontal cortex including the orbitobasal surface (BA 47/11).

[Please place Fig. 3 near here]

Connectivity within the Prosody Network

So far, three successive steps of prosody processing have been identified: (1) extraction of supra-segmental acoustic information, (2) representation of supra-segmental sequences and (3) explicit judgment of emotional information. As concerns the respective neuroanatomical correlates, extraction of suprasegmental acoustic information seems to be predominantly bound to the right primary and secondary auditory regions. Presumably, the relevant acoustic information is transferred from these regions via direct fibre connections to an area within the posterior superior temporal sulcus (post-STS) subserving the representation of meaningful intonational sequences. In case of explicit judgement of emotional prosody, a further temporo-frontal passage of information must be assumed accounting for the observed activation of bilateral inferior-frontal cortex during this task. It should be emphasized, furthermore, that converging results from lesion studies (Hornak et al., 1996, 2003; Ross et al., 1997) and functional imaging examinations (Imaizumi et al., 1997; Pihan et al., 2000; Wildgruber et al., 2002, 2004) suggest a contribution of these areas to the processing of emotional prosody, and an intact transcallosal communication of information has been assumed to be a prerequisite for comprehension of emotional prosody (Ross et al., 1997). It is unclear, however, whether this cooperation of the two hemispheres is based on a sequence of processing steps or if both frontal lobes receive the respective information independently via parallel connections from the right posterior temporal cortex. In order to investigate the connectivity architecture of the cerebral network involved in the processing of emotional prosody, a further experiment was carried out. Twenty-four healthy right handed subjects (11 males, 13 females, mean age 24.4 years)

underwent event-related fMRI measurements while rating the emotional valence of either prosody or semantics of 162 binaurally presented emotional adjectives (54 x neutral, 54 x positive, 54 x negative content) spoken in happy, neutral or angry intonation by six professional actors (3 females / 3 males). The adjectives were selected from a sample of 500 adjectives on the basis of ratings obtained from 45 healthy German native speakers (see Kissler et al, this volume) along the dimensions of valence and arousal on a 9-point self assessment manikin scale (SAM, Bradley and Lang, 1994). The stimuli comprised 54 highly arousing positive (mean arousal rating > 4, mean valence rating < 4, e.g. “verführerisch” = alluring), 54 highly arousing negative (mean arousal > 4, mean valence rating > 6, e.g., “panisch” = panic) and 54 low arousing neutral (mean arousal rating < 4, mean valence rating between 4 and 6, e.g. “breit” = broad). During separate functional imaging sessions, subjects had been asked to judge either the valence of emotional word content or the valence of emotional prosody on the 9 point SAM-scale. Both the order of within-session stimulus presentation and the sequence of sessions were pseudo-randomized across subjects. To assess functional connectivity of activated regions, the novel technique of dynamic causal modelling (Friston et al., 2003) was applied to the data. This approach allows inferences on (1) the parameters representing influence of experimentally designed inputs, (2) the intrinsic coupling of different brain regions, and (3) modulation of this coupling by experimental factors (for methodological details see Ethofer et al, 2006). Using this technique, the following hypotheses were evaluated:

- a) Within the network of regions characterized by task-dependent activation, the post-STS serves as input region (receiving input from primary and secondary acoustic regions).
- b) The frontal lobes, consecutively, receive their input from the post-STS. Moreover, it was assessed whether both frontal lobes subserve two successive processing steps or receive their information independently from the right post-STS via parallel pathways.

Conventional analysis of the fMRI data yielded – in very good accordance with prior investigations (Wildgruber et al., 2004, 2005) - activation within the right posterior STS and bilateral inferior frontal cortices during evaluation of emotional prosody. Subsequent determination of functional connectivity revealed that the activation cluster within the right post-STS represents the most likely input region into this task-specific network. This finding is in agreement with the assumption that this region subserves representation of suprasegmental sequences and receives direct input from primary and secondary acoustic regions. To investigate the intrinsic connectivity pattern within the network, dynamic causal models assuming parallel, serial or fully bidirectional connectivity patterns were compared. The model based upon parallel projections from the posterior STS to the frontal cortical regions turned out to be significantly superior to both serial models as well as the model with bilaterally connected brain regions (Fig. 4a). In a post hoc analysis, an attempt was made to optimize this parallel pathway model by adding either unidirectional or bidirectional connections between the two frontal regions or adding unilateral or bidirectional backward projections from the frontal areas to the right posterior STS. The original parallel pathway model again was found to be significantly superior to all

alternative models. These results provide further empirical support for the hypothesis, that processing of emotional prosody is carried out in three successive steps: (1) extraction of suprasegmental acoustic information bound to predominantly right-sided primary and higher-order acoustic regions, (2) representation of meaningful suprasegmental sequences within the right post-STS and (3) explicit emotional judgment of acoustic information within the bilateral inferior frontal cortices (Fig. 4b).

[Please place Fig. 4 near here]

Implicit Processing of Emotional Prosody

During every-day interactions among humans, as a rule, the emotional connotations of communicative signals are not explicitly evaluated on a quantitative scale. Rather, highly-automatized understanding of the emotional information conveyed by facial expressions, speech prosody, gestures or the propositional content of verbal utterances seems to be much more important. A variety of empirical data indicate different cerebral pathways to be involved in explicit and implicit processing of emotional signals (LeDoux et al., 1996; Anderson and Phelps, 1998; Adolphs, 1999; Critchley, 2000; Adolphs et al., 2002). As concerns the hemodynamic responses bound to specific emotional categories, a selective contribution of the amygdala to recognition of fearful voices has been assumed on the basis of lesion data (Scott et al., 1997) and prior PET studies (Phillips et al. 1998; Morris et al., 1999). Furthermore, a specific contribution of the anterior insula and the basal ganglia to the perception of vocal expressions of disgust has been predicted based

on clinical findings (Pell and Leonhard, 2003) and functional imaging experiments during processing of facial expressions (Sprengelmeyer et al., 1998; Pihan et al., 2002; Wicker et al., 2003). Responses of the amygdalae have been observed to depend on implicit processing of emotional signals, e.g., during passive listening tasks, whereas explicit judgments of emotional expressions were shown to result in deactivation of this region (Morris et al., 1999; Critchley et al., 2000; Adolphs 2002). As a consequence, implicit transmission of emotional information by the induction of physiological emotional reactions, e.g., changes of heart rate and skin conductance, might be linked to emotion-specific subcortical regions whereas the explicit evaluation of emotional signals based on the retrieval of information from emotional memory appears to be processed within bilateral inferior-frontal areas, irrespective of emotion type and valence of the stimuli. In order to evaluate the neural basis of implicit processing of emotional prosody, a cross-modal interaction experiment was conducted (for methodological issues of cross-modal interaction experiments see Ethofer et al., this volume). This experiment was designed to test the following two predictions:

- a) Simultaneous presentation of emotional faces and emotional prosody induces distinct interaction effects: explicit judgment of facial expressions is influenced by implicit processing of unattended emotional prosody.
- b) The impact of an unattended fearful tone of speech on explicit judgment of emotional faces is associated with activation of the amygdala.

During this experiment, images of facial expressions taken from the Ekman and Friesen battery (Ekman and Friesen, 1976) were presented to 12 healthy right-

handed subjects (7 males, 5 females, age: 19-29 years). Using digital morphing techniques, a series of visual stimuli was generated extending in facial expression from 100% fear to 100% happiness in incremental steps of 25% (Perret et al., 1994). In one run of the experiment, the facial expressions were shown in isolation, and in another trial they were combined with acoustic stimuli, i.e., short declarative sentences spoken in a fearful or happy tone by two professional actors (one male, one female). In both of these runs participants were instructed to rate the emotional valence of the displayed facial expressions. A third run of the experiment required explicit judgment of emotional prosody. The behavioral results show that subjects rated fearful and neutral facial expressions as being more fearful when presented concomitant with a fearfully spoken sentence as compared to the no-voice condition. By contrast, no significant shifts in interpretation occurred during application of happy expressions (Fig. 5a). Thus, this experimental paradigm might provide a means for quantitative measurements of the implicit impact of emotional prosody on the judgment of facial expressions (de Gelder and Vroomen, 2000). A comparison of happy and fearful intonations during explicit judgment of prosody (unimodal auditory session) did not reveal any significant differences of the hemodynamic cerebral responses. As concerns implicit processing of emotional prosody, however, the middle section of the right fusiform gyrus showed a significantly stronger activation when facial expressions were displayed in the presence of a fearful voice as compared to happy intonation. This region has been named the fusiform face area, because it has been found crucial for the processing of faces in clinical and experimental studies (Barton et al. 2000, Puce et al. 1995, Kanwisher et al. 1997). Moreover, this region shows stronger activation to emotional as compared to neutral faces (Morris et al. 1998) and seems to respond particularly to stimuli signalling

danger (Surguladze et al., 2003). The increased hemodynamic responses within the fusiform gyrus in presence of an auditory expression of threat might reflect enhanced alertness for detection of the respective visual cues, giving rise to shifts in the interpretation of facial expressions. Moreover, comparison of hemodynamic responses with the individual explicit ratings of emotional facial expressions in presence of unattended fearful prosody revealed a significant correlation within the basolateral part of the left amygdala extending into the periamygdaloid cortex. This finding indicates the impact of voice on the processing of faces to be mediated via anterior-temporal structures. In line with this assumption, the amygdala has been observed to modulate neuronal activity in brain regions subserving visual processing (Morris et al. 1998, Davis and Whalen 2001, Vuilleumier et al. 2004), and it has been suggested the left-sided nuclei integrate audiovisual fear-related emotional information into a common percept (Dolan et al. 2001).

[Please place Fig. 5 near here]

Cross-modal integration of emotional communicative signals

Emotional information may be conveyed via different communicative channels, e.g., prosodic features of the acoustic speech signal, facial expressions, and propositional content of verbal utterances. Based on the findings presented here, several successive steps during cross-modal integration of emotional signals can be separated and assigned to distinct cerebral correlates: (1) extraction of communicative signals is subserved by the respective modality-specific primary

cortices, (2) modality-specific higher-order auditory regions process emotional information (i.g., prosody = right STS, facial expressions = fusiform face area, propositional meaning = left posterior STG), (3) explicit emotional judgments, presumably involving evaluation of associations with episodic emotional memory, were found to be linked to bilateral orbitofrontal cortex. Implicit processing of emotional signals, however, seems to rely on alternative pathways including emotion-specific subcortical regions involved in automatic physiological reaction (e.g. variation of heart rate and skin conductance). It has been demonstrated, that both pathways of emotion processing influence the behavior of the organism and that unattended processing of emotional information may interact with attended evaluation of emotional communicational signals (Fig. 5b). Future research will be required, however, to further clarify the neuroanatomical basis of interaction effects between implicit and explicit stimulus processing and integration of emotional signals conveyed by various means of communication.

Acknowledgements:

The reported studies were supported by the 'Junior science program of the Heidelberger Academy of Sciences and Humanities' and the German Research Foundation (DFG WI 2101 and SFB 550 B10).

Abbreviations:

BA:	Brodmann Area
fMRI:	functional magnetic resonance imaging
IFC:	inferior frontal cortex
IFG:	inferior frontal gyrus
mid-STG:	middle part of the superior temporal sulcus
MTG:	middle temporal gyrus
post-STG:	posterior part of the superior temporal sulcus
STG:	superior temporal gyrus
STS:	superior temporal sulcus

References:

- Ackermann H, Hertrich I, Ziegler W (1993) Prosodische Störungen bei neurologischen Erkrankungen: Eine Literaturübersicht. *Fortschr. Neurol. Psychiat.*, 61: 241-253.
- Ackermann H, Riecker A, Grodd W, Wildgruber D (2001) Rate-dependent activation of a prefrontal-insular-cerebellar network during passive listening to trains of click stimuli: An fMRI study, *NeuroReport*, 18: 4087-4092.
- Ackermann H, Hertrich I, Grodd W, Wildgruber D (2004) „Das Hören von Gefühlen“: Funktionell-neuroanatomische Grundlagen der Verarbeitung affektiver Prosodie. *Aktuelle Neurologie*, 31: 449-460.
- Ackermann H, Riecker A, Wildgruber D (2006) Cerebral correlates of singing capabilities in humans: Clinical observations, experimental-behavioural studies, and functional imaging data. In: Altenmüller E, Kesselring J, Wiesendanger M (Eds.). *Music, Motor Control, and the Brain*. Oxford University Press, Oxford, in Press
- Adolphs R, Tranel D (1999) Intact recognition of emotional prosody following amygdala damage. *Neuropsychologia*, 37: 1285-1292.
- Adolphs R, Tranel D, Damasio H (2001) Emotion recognition from faces and prosody following temporal lobectomy. *Neuropsychology*, 15: 396-404.
- Adolphs R (2002) Neural systems for recognizing emotion. *Curr. Opin. Neurobiol.*, 12: 169-177.
- Adolphs R, Damasio H, Tranel D (2002) Neural systems for recognition of emotional prosody: a 3-D lesion study. *Emotion*, 2: 23-51.
- Anderson AK, Phelps EA (1998) Intact recognition of vocal expressions of fear following bilateral lesion of the human amygdala. *Neuroreport*, 9: 3607-3613.

- Bachorowski JO, Owren MJ (2003) Sounds of Emotion: Production and Perception of Affect-Related Vocal Acoustics. *Ann. NY Acad. Sci.*, 1000: 244-265.
- Banse R, Scherer KR (1996) Acoustic Profiles in Vocal Emotion Expression. *Journal of Personality and Social Psychology*, 70: 614-636.
- Barton JJS, Press DZ, Keenan JP, O'Connor M (2002) Lesion of the fusiform face area impair perception of facial configuration in prosopagnosia. *Neurology*, 58: 71-78.
- Baum SR, Pell MD (1999) The neural basis of prosody: insights from lesion studies and neuroimaging. *Aphasiology*, 13: 581-608.
- Behrens SJ (1985) The perception of stress and lateralization of prosody. *Brain Lang.*, 26: 332-348.
- Belin P, Zilbovicius M, Crozier S, Thivard L, Fontaine A, Masure MC, Samson Y (1998) Lateralization of speech and auditory temporal processing. *Journal of Cognitive Neuroscience*, 10: 536-540.
- Blair RJR, Morris JS, Frith CD, Perret DI, Dolan RJ (1999) Dissociable neural responses to facial expressions of sadness and anger. *Brain* 122: 883-893.
- Blair RJR, Cipolotti L (2000) Impaired social response reversal. *Brain*, 123: 1122-1141.
- Borod JC, Obler LK, Erhan HM, Grunwald IS, Cicero BA, Welkowitz J, Santschi C, Agosti RM, Whalen JR (1998) Right Hemisphere Emotional Perception: Evidence Across Multiple Channels. *Neuropsychology*, 12: 446-458.
- Borod JC, Zgaljardic D, Tabert MH, Koff E (2001) Asymmetries of emotional perception and expression in normal adults. In: Gainotti G (Ed.) *Handbook of Neuropsychology*, 2nd Edition, Vol. 5, Elsevier Science, 181-205.

- Borod JC, Bloom RL, Brickman AM, Nakhutina L, Curko EA (2002) Emotional Processing Deficits in Individuals With Unilateral Brain Damage. *Applied Neuropsychology*, 9: 23-36.
- Bradley MM, Lang PJ (1994) Measuring emotion: The Self-Assessment Manikin and the Semantic Differential. *J Behav Ther Exp Psychiatry*, 25: 49-59.
- Breitenstein C, Daum I, Ackermann H (1998) Emotional processing following cortical and subcortical brain damage: Contribution of the fronto-striatal circuitry. *Behavioural Neurology*, 11: 29-42.
- Breitenstein C, Van Lancker D, Daum I, Waters CH (2001) Impaired Perception of Vocal Emotions in Parkinson's Disease: Influence of Speech Time Processing and Executive Functioning. *Brain and Cognition*, 45: 277-314.
- Buchanan TW, Lutz K, Mirzazade S, Specht K, Shah NJ, Zilles K, Jäncke L (2000) Recognition of emotional prosody and verbal components of spoken language: an fMRI study. *Cognitive Brain Research*, 9: 227-238.
- Calder AJ, Keane J, Manes F, Antoun N, Young AW (2000) Impaired recognition and experience of disgust following brain injury. *Nat Neurosci*, 3: 1077-1078.
- Cancelliere AE, Kertesz A (1990) Lesion Localization in Acquired Deficits of Emotional Expression and Comprehension. *Brain and Cognition*, 13: 133-147.
- Canli T, Desmond JE, Zhao Z, Glover G, Gabrieli JD (1998) Hemispheric asymmetry for the emotional stimuli detected with fMRI. *Neuroreport*, 9: 3233-3239.
- Charbonneau S, Scherzer BP, Aspirot D, Cohen H (2003) Perception and production of facial and prosodic emotions by chronic CVA patients. *Neuropsychologia*, 41: 605-613.
- 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. *Human Brain Mapping*, 9: 93-105.

Cutler A, Dahan D, Donselaar W (1997) Prosody in the Comprehension of Spoken Language: A Literature Review. *Language and Speech*, 40: 141-201.

Darby DG (1993) Sensory aprosodia: a clinical clue to lesions of the inferior division of the right middle cerebral artery? *Neurology*, 43: 567-72.

Davidson RJ, Abercrombie H, Nitschke JB, Putnam K (1999) Regional brain function, emotion and disorders of emotion. *Curr Opin Neurobiol*, 9: 228-234.

Davis M, Whalen PJ (2001): The amygdala: Vigilance and emotion. *Mol Psychiatry*, 6: 13-34.

De Gelder B, Vroomen J (2000) The perception of emotions by ear and eye. *Cognition and Emotion*, 14: 289-311.

Dolan RJ, Morris JS, De Gelder B (2001) Crossmodal binding of fear in voice and face. *Proc Natl Acad Sci USA*, 98: 10006-10010.

Ekman P, Friesen W (1976) *Pictures of facial affect*, Consulting Psychologists Press, Palo Alto.

Emmorey KD (1987) The neurological substrates for prosodic aspects of speech. *Brain Lang*, 30: 305-329.

Ethofer T, Anders S, Erb M, Droll C, Royen L, Saur R, Reiterer S, Grodd W, Wildgruber D (2006a) Impact of voice on emotional judgement of faces: an event-related fMRI study, *Human Brain Mapping*, in press.

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*, in press.

- Ethofer T, Erb M, Anders S, Wiethoff S, Herbert C, Saur R, Grodd W, Wildgruber D (2006c) Effects of prosodic emotional intensity on activation of associative auditory cortex, *NeuroReport*, in press
- Friston KJ, Harrison L, Penny W (2003) Dynamic causal modeling. *NeuroImage*, 19: 1273-1302.
- Gandour J, Wong D, Hutchins G (1998) Pitch processing in the human brain is influenced by language experience. *NeuroReport*, 9: 2115-2119.
- Geigenberger A, Ziegler W (2001) Receptive prosodic processing in aphasia. *Aphasiology*, 15: 1169-1188.
- George MS, Parekh PI, Rosinsky N, Ketter TA, Kimbrell TA, Heilman KM, Herscovitch P, Post RM (1996) Understanding emotional prosody activates right hemisphere regions. *Arch Neurol*, 53: 665-670.
- Grandjean D, Sander D, Pourtois G, Schwartz S, Seghier ML, Scherer KR, Vuilleumier P (2005) The voices of wrath: brain responses to angry prosody in meaningless speech. *Nat Neurosci*, 8: 145-146.
- Heilman KM, Scholes R, Watson RT (1975) Auditory affective agnosia: disturbed comprehension of affective speech. *J Neurol Neurosurg Psychiatry* 38: 69-72.
- Heilman KM, Bowers D, Speedie L, Coslett HB (1984) Comprehension of affective and nonaffective prosody. *Neurology*, 34: 917-921.
- Hornak J, Rolls ET, Wade D (1996) Face and voice expression identification in patients with emotional and behavioral changes following ventral frontal lobe damage. *Neuropsychologia*, 34: 247-261.
- Hornak J, Bramham J, Rolls ET, Morris RG, O'Doherty J, Bullock PR, Polkey CE (2003) Changes in emotion after circumscribed surgical lesions of the orbitofrontal and cingulate cortices. *Brain*, 126: 1691-1712.

- Hugdahl K, Davidson RJ (2003) *The Asymmetrical Brain*. Cambridge, London, MIT Press
- Hughling-Jackson (1879) On affections of speech from disease of the brain (reprint from *Brain* 1879). *Brain* (1915), 38: 107-129.
- Imaizumi S, Mori K, Kiritani S, Kawashima R, Sugiura M, Fukuda H, Itoh K, Kato T, Nakamura A, Hatano K, Kojima S, Nakamura K (1997) Vocal identification of speaker and emotion activates different brain regions. *NeuroReport*, 8: 2809-2812.
- Ivry RB, Robertson LC (1998) *The Two Sides of Perception*. Cambridge, MA, MIT Press.
- Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci*, 17: 4302-4311.
- Kesler-West ML, Andersen AH, Smith CD, Avison MJ, Davis CE, Kryscio RJ, Blonder LX (2001) Neural substrates of facial emotion processing using fMRI. *Cogn. Brain Res.*, 11: 213-226.
- Kotz SA, Meyer M, Alter K, Besson M, von Cramon DY, Friederici AD (2003) On the lateralization of emotional prosody: an event-related functional MR investigation. *Brain Lang*, 86: 366-376.
- Kucharska-Pietura K, Phillips ML, Gernand W, David AS (2003) Perception of emotions from faces and voices following unilateral brain damage. *Neuropsychologia*, 41: 1082-1090.
- LeDoux J (1996) *The emotional brain*, New York, Simon and Schuster.
- Lehiste I (1970) *Suprasegmentals*, MIT Press, Cambridge, MA.

Levin HS, Eisenberg HM, Benton AL (1991) *Frontal Lobe Function and Dysfunction*. Oxford University Press, New York, S. 318-338.

Mehrabian A (1972) *Nonverbal communication*, Chicago, Albine-Atherton.

Menon RS, Goodyear BG (1999) Submillimeter functional localization in human striate cortex using BOLD contrast at 4 Tesla: implications for the vascular point-spread function. *Magn Reson Med*, 41:230-235.

Meyer M, Alter K, Friederici AD, Lohmann G, von Cramon DY (2002) FMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. *Human Brain Mapping*, 17: 73-88.

Mitchell RLC, Elliot R, Barry M, Cruttenden A, Woodruff PWR (2003) The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. *Neuropsychologia*, 41: 1410-1421.

Morris JS, Frith CD, Perrett DI, Rowland D, Young AW, Calder AJ, Dolan RJ (1996): A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature*, 383: 812-815.

Morris JS, Friston KJ, Büchel C, Frith CD, Young AW, Calder AJ, Dolan RJ (1998): A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain*, 121: 47-57.

Morris JS, Scott SK, Dolan RJ. (1999) Saying it with feelings: neural responses to emotional vocalizations. *Neuropsychologia*, 1999; 37: 1155-1163.

Murphy FC, Nimmo-Smith I, Lawrence AD (2003) Functional neuroanatomy of emotions: a meta-analysis. *Cogn Affect Behav Neurosci*, 3: 207-233.

Murray IR, Arnott JL (1993) Toward the simulation of emotion in synthetic speech: a review of the literature on human vocal emotion. *J Acoust Soc Am*, 93: 1097-1108.

- Nakamura K, Kawashima R, Ito K, Sugiura M, Kato T, Nakamura A, Hatano K, Nagumo S, Kubota K, Fukuda H, Kojima S (1999) Activation of the right inferior frontal cortex during assessment of facial emotion. *J Neurophysiol*, 82: 1610-1614.
- Pell MD, Baum SR (1997) The ability to perceive and comprehend intonation in linguistic and affective contexts by brain-damaged adults. *Brain Lang*, 57: 80-99.
- Pell MD, Baum SR (1997) Unilateral brain damage, prosodic comprehension deficits, and the acoustic cues to prosody. *Brain Lang*, 57: 195-214.
- Pell MD (1998) Recognition of prosody following unilateral brain lesions: influence of functional and structural attributes of prosodic contours. *Neuropsychologia*, 36: 701-715.
- Pell MD, Leonard CL (2003) Processing emotional tone from speech in Parkinson's disease: A role for the basal ganglia. *Cognitive, Affective, & Behavioral Neuroscience*, 3: 275-288.
- Peper M, Irle E (1997) Categorical and dimensional decoding of emotional intonations in patients with focal brain lesions. *Brain Lang*, 58: 233-264.
- Perrett I, May KA, Yoshikawa S (1994): Facial shape and judgements of female attractiveness. *Nature*, 368:239-42.
- Phan KL, Wager T, Taylor SF, Liberzon I (2002) Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage*, 16: 331-348.
- Phillips ML, Young AW, Scott SK, Calder AJ, Andrew C, Giampietro V, Williams SCR, Bullmore ET, Brammer M, Gray JA (1998). Neural responses to facial and vocal expressions of fear and disgust. *Proc R Soc Lon*, 265: 1809-1817.

- Pihan H, Altenmüller E, Ackermann H (1997) The cortical processing of perceived emotion: a DC-potential study on affective speech prosody. *NeuroReport*, 8: 623-627.
- Pihan H, Altenmüller E, Hertrich I, Ackermann H (2000) Cortical activation patterns of affective speech processing depend on concurrent demands on the subvocal rehearsal system: A DC-potential study. *Brain*, 123: 2338-2349.
- Pitkänen A (2000) *The Amygdala. A functional analysis*, Oxford University Press, New York.
- Poeppel D, Guillemin A, Thompson J, Fritz J, Bavelier D, Braun A (2004) Auditory lexical decision, categorical perception, and FM direction discrimination differentially engage left and right auditory cortex. *Neuropsychologia*, 42: 183-200.
- Price JL (1999) Prefrontal Cortical Network Related to Visceral Function and Mood. *Ann N Y Acad Sci*, 877: 383-396.
- Puce A, Allison T, Gore J, McCarthy G (1995): Face-sensitive regions in human extrastriate cortex studied by functional MRI. *J Neurophysiol*, 74: 1192-1199.
- Reiterer SM, Erb M, Droll CD, Anders S, Ethofer T, Grodd W, Wildgruber D (2005) Impact of task difficulty on lateralization of pitch and duration discrimination. *NeuroReport*, 16: 239-242
- Rolls ET (1999) The functions of the orbito-frontal cortex. *Neurocase*, 5: 301-312.
- Ross ED (1981) The aprosodias: functional-anatomic organization of the affective components of language in the right hemisphere. *Arch Neurol*, 38: 561-569.
- Ross ED, Thompson RD, Yenkosky J (1997) Lateralization of affective prosody in the brain and the callosal integration of hemispheric language functions. *Brain Lang*, 56: 27-54.

- Scherer KR, Johnstone T, Klasmeyer G (2003) Vocal expression of emotion. In: Davidson RJ, Scherer KR, Goldsmith HH (Eds.) *Handbook of Affective Sciences*, New York, Oxford, S. 433-456.
- Schirmer A, Alter K, Kotz S, Friederici AD (2001) Lateralization of prosody during language production: a lesion study. *Brain Lang*, 76: 1-17.
- Schmitt JJ, Hartje W, Williams K (1997) Hemispheric asymmetry in the recognition of conditional attitude conveyed by facial expression, prosody and propositional speech. *Cortex*, 33: 65-81.
- Scott SK, Young AW, Calder AJ, Hellowell DJ (1997) Impaired auditory recognition of fear and anger following bilateral amygdala lesions. *Nature*, 385: 254-275.
- Sidtis JJ, Van-Lancker-Sidtis D (2003) A neurobehavioral approach to dysprosody. *Semin Speech Lang*, 24: 93-105.
- Small DM, Zatorre RJ, Dagher A, Evans AC, Jones-Gotman M (2001) Changes in brain activity related to eating chocolate: From pleasure to aversion. *Brain*, 124: 1720-1733.
- Sprengelmeyer R, Rausch M, Eysel UT, Przuntek H (1998) Neural structures associated with recognition of facial expressions of basic emotions. *Proc R Soc Lond B Biol Sci*, 265: 1927-1931.
- Starkstein SE, Federoff JP, Price TR, Leiguarda RC, Robinson RG (1994) Neuropsychological and neuroradiologic correlates of emotional prosody comprehension. *Neurology*, 44: 515-522.
- Surguladze SA, Brammer MJ, Young AW, Andrew C, Travis MJ, Williams SCR, Phillips ML (2003) A preferential increase in the extrastriate response to signals of danger. *NeuroImage*, 19: 1317-1328.

- Van Lancker D (1980) Cerebral lateralization of pitch cues in the linguistic signal. *Int J Human Commun*, 13: 227-277.
- Van Lancker D, Sidtis JJ (1992) The identification of affective-prosodic stimuli by left- and right-hemisphere-damaged subjects: all errors are not created equal. *J Speech Hear Res*, 35: 963-970.
- Vuilleumier P, Richardson MP, Armony JL, Driver J, Dolan RJ (2004): Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nat Neurosci*, 7: 1271-1278.
- Weddell R (1994) Effects of subcortical lesion site on human emotional behaviour. *Brain and Cognition*, 25: 161-193.
- Wicker B, Keysers C, Plailly J, Royet JP, Gallese V, Rizzolatti G (2003) Both of us disgusted in My insula: the common neural basis of seeing and feeling disgust. *Neuron*, 40: 655-664.
- Wildgruber D, Ackermann H, Klose U, Kardatzki B, Grodd W (1996) Functional Lateralization of Speech Production at Primary Motor Cortex: A fMRI Study. *NeuroReport*, 7: 2791-2795.
- Wildgruber D, Erb M, Klose U, Grodd W (1997) Sequential activation of supplementary motor area and primary motor cortex during self-paced finger movement in human evaluated by functional MRI. *Neuroscience Letters*, 127: 161-164.
- Wildgruber D, Ackermann H, Klose U, Kardatzki B, Grodd W (1998) Hemispheric Lateralization of Speech Production and Singing at the Level of the Motor Cortex in fMRI. In: Ziegler W, Deger K (Eds.) *Clinical Phonetics and Linguistics*, Whurr, London, S. 238-243.

- Wildgruber D, Kischka U, Faßbender K, Ettl T (2000) The Frontal Lobe Score: Evaluation of its clinical validity. *Clinical Rehabilitation*, 14: 272-278.
- Wildgruber D, Ackermann H, Grodd W (2001) Differential contributions of Motor cortex, basal ganglia and Cerebellum to Speech Motor control: Effects of syllable repetition rate evaluated by fMRI. *NeuroImage*, 13: 101-109.
- Wildgruber D, Pihan H, Ackermann H, Erb M, Grodd W (2002) Dynamic Brain Activation during Processing of Emotional Intonation: Influence of Acoustic Parameters, Emotional Valence and Sex. *NeuroImage*, 15: 856-869.
- Wildgruber D, Ackermann H (2003) Aphasie. In: Brandt T, Dichgans J, Diener HC (Eds.) *Therapie und Verlauf neurologischer Erkrankungen*, Kohlhammer, Stuttgart, 267-277.
- 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.
- Wildgruber D, Riecker A, Hertrich I, Erb M, Grodd W, Ethofer T, Ackermann H (2005) Identification of Emotional Intonation evaluated by fMRI. *NeuroImage*, 24: 1233-1241.
- Zatorre RJ (2001) Neural specializations for tonal processing. *Ann NY Acad Sci*, 930: 193-210.
- Zatorre RJ, Belin P (2001) Spectral and Temporal Processing in human auditory cortex. *Cerebral Cortex*, 11: 946-953.
- Zatorre RJ, Belin P, Penhune V (2002) Structure and Function of Auditory Cortex: Music and Speech, *Trends in Cognitive Science*, 6: 37-46.

Figure Legends:

Fig. 1: (a) According to the acoustic lateralization hypothesis, rapid changes of acoustic parameters (< 50 ms) are predominantly processed within the left hemisphere, whereas slow variations (> 100 ms) of these parameters are mainly processed within the right hemisphere.

(b) Parametric responses during passive listening to trains of noise bursts: hemodynamic responses characterized by positive linear (red), negative linear (green) or non-linear (blue) rate-response functions. Activation clusters are displayed on transverse sections of the averaged anatomical reference images (R = right, L = left). The relationship between signal intensity and rate of acoustic stimulation (in arbitrary units) was determined within the right (green) and left (blue) insular cortex (see Ackerman et al., 2001).

(c) Discrimination of sound duration: pairs of complex acoustic signals that varied in duration (100 to 400 ms) were presented to healthy subjects. Accuracy rates demonstrate increasing deviance in time to be correlated with higher performance scores. Parametric effects: significantly activated areas as a function of linear increase with task performance emerged within the right MTG/STG and the left IFG during duration discrimination. Laterality analysis: voxelwise comparison of the hemispheres revealed a significantly activated cluster within the left STG for the parametric effect of duration discrimination (see Reiterer et al., 2005).

(d) Parametric effects of prosodic emotional intensity. Conjunction of regions showing a linear relationship between hemodynamic responses and prosodic emotional intensity during both implicit and explicit processing of emotional prosody. Beta estimates (mean \pm standard error) corresponding to distinct intensity steps of

emotional intonations have been plotted for the most significant voxel of the cluster in the right and left STS during implicit (red) and explicit (green) processing of emotional prosody (see Ethofer et al., 2006c).

Fig. 2: (a) According to the valence hypothesis positive emotional information (i.e. happy expressions) is processed within the left hemisphere, whereas negative emotional information (expressions of fear, anger, disgust or sadness) is processed within the right hemisphere.

(b) Significant hemodynamic responses during identification of emotional intonation as compared to vowel identification are superimposed upon the cortical surface of a template brain and upon an axial slice at the level of the highest activated voxels within the activation clusters. The emotional task yielded specific activation within the right STS (BA 22/42) and the right inferior frontal cortex (BA 45/47). Analysis of valence effects, however, revealed no differences of cerebral responses depending upon valence or specific emotional categories (see Wildgruber et al., 2005).

Fig. 3: (a) According to the functional lateralization hypothesis linguistic prosody is processed within the left hemisphere, whereas emotional prosody is bound to the right hemisphere.

(b) Variation of linguistic (left) and emotional intonation (right). The German sentence “Der Schal ist in der Truhe” (“the scarf is in the chest”) was digitally resynthesized with various pitch contours. Five different patterns of sentence focus were realised by a stepwise increase of the fundamental frequency on the final word (left). The stress accentuation ranged between an utterance clearly focussed on the second word (solid line) and one that is focussed on the final word (dotted line). For each of these

synthetic sentences, five variations of emotional expressiveness were generated by manipulation of the pitch range across the whole utterance (right). Sentences with broader pitch ranges are perceived as being more excited. As shown for the middle contour (red) the realisation of linguistic accents remains constant during manipulation of emotional expressiveness. The sentences of each stimulus pair differed in relative focus accentuation as well as in emotional intensity.

(c) Significantly activated regions, identified by task comparisons, superimposed upon the cortical surface of a template brain and upon an axial slice at the level of the highest activated voxels within each activation cluster. The emotional task (upper row) yielded significant responses within the bilateral orbitobasal frontal cortex (BA 11/47), whereas activation of the left inferior frontal gyrus (BA 44/45) emerged during discrimination of linguistic prosody (lower row) (see Wildgruber et al., 2004).

Fig 4: (a) To evaluate the intrinsic connectivity of regions contributing to processing of emotional prosody four different models were compared. (Model 1) Parallel transmission from the right post-STS to both frontal regions. (Model 2) Successive conductance from post-STS to right IFG and further on to the left IFG. (Model 3) Serial conductance from post-STS to left IFG and right IFG. (Model A) Fully connected bidirectional flow of information. Based upon a prior analysis, in all these models external inputs were specified to enter the network via the right post-STS. Dynamic causal modelling revealed a statistical superiority of the parallel processing model (Model 1) as compared to all other models (Ethofer et al., 2006b).

(b) Based on these findings it is assumed that explicit judgment of emotional prosody is carried out in at least three successive steps: (1) extraction of suprasegmental information bound to predominantly right-sided primary and

secondary acoustic regions, (2) representation of meaningful suprasegmental sequences within the right post-STS and (3) explicit emotional judgment of acoustic information within the bilateral inferior frontal cortices.

Fig. 5: (a) Implicit impact of fearful prosody on judgment of emotional faces: (left) valence rating of facial expressions (mean \pm standard error) presented without acoustic stimuli (white bars) and in combination with fearful prosody (grey). Evaluation of facial expressions in presence of a fearful voice as compared to a happy intonation yielded significant activation in the right fusiform gyrus (upper right). Analysis of cross-modal impact of fearful voices revealed significant correlations between individual behavioural changes and hemodynamic responses in the left amygdala (see Ethofer et al., 2006a).

(b) Cross-modal integration of emotional communicative signals: (1) extraction of different communicative signals (prosody, facial expressions, word content) is subserved by the respective modality-specific primary cortices. (2) More complex features of these signals are processed within modality specific secondary regions. (3) As a third step, explicit emotional judgments based on evaluation of associations with episodic emotional memory seem to be linked to the bilateral inferior frontal cortex. This region is assumed to be involved in cross-modal integration during explicit evaluation. On the other hand, emotional signals can yield an automatic (implicit) induction of emotional physiological reaction (e.g. variation of heart rate and skin conductance) that is linked to specific subcortical regions. Presumably, both neural pathways are inter-connected at various levels.

Fig. 1)

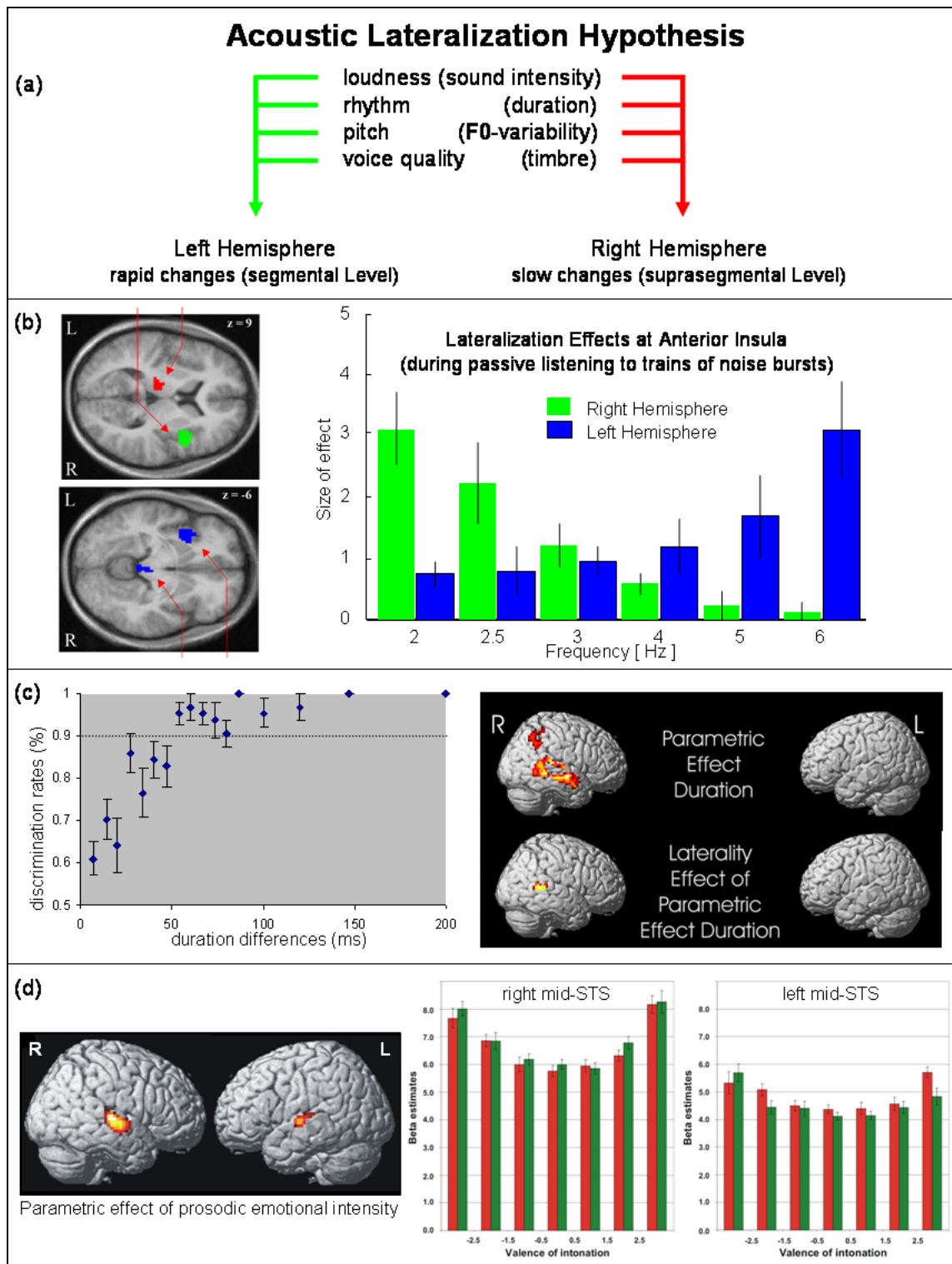


Fig. 2)

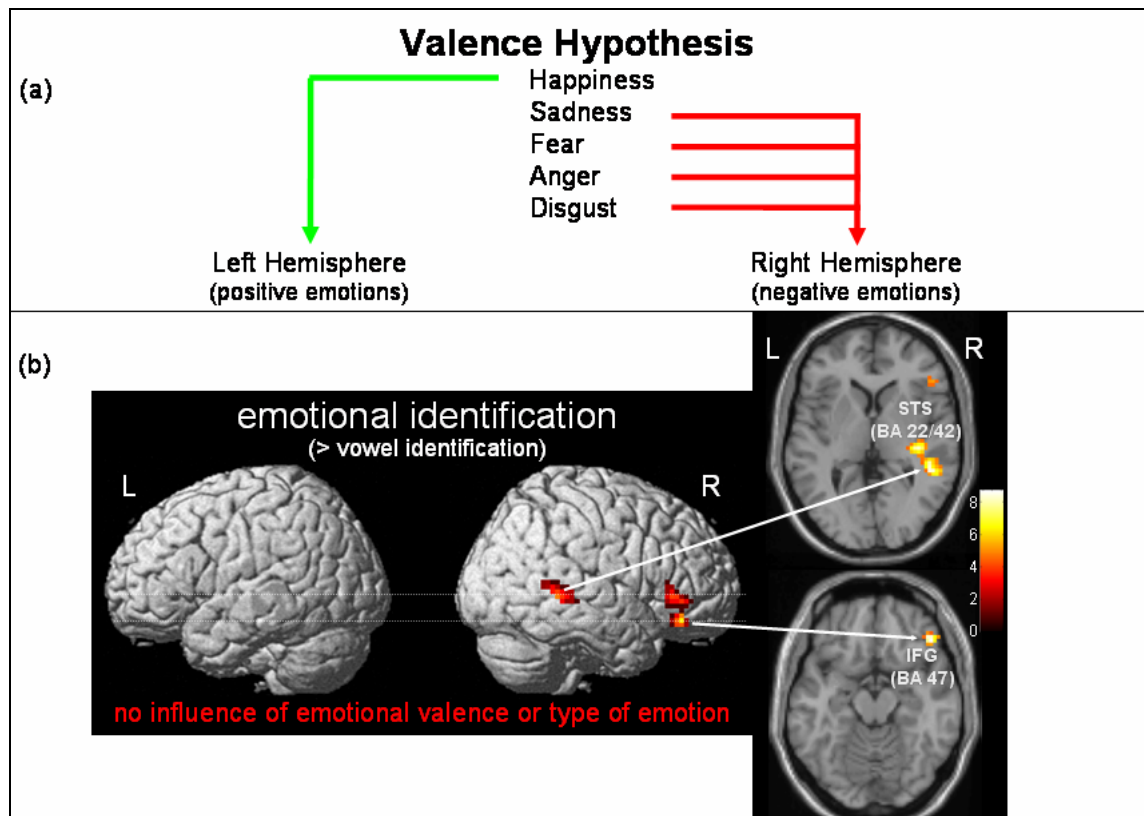


Fig. 3)

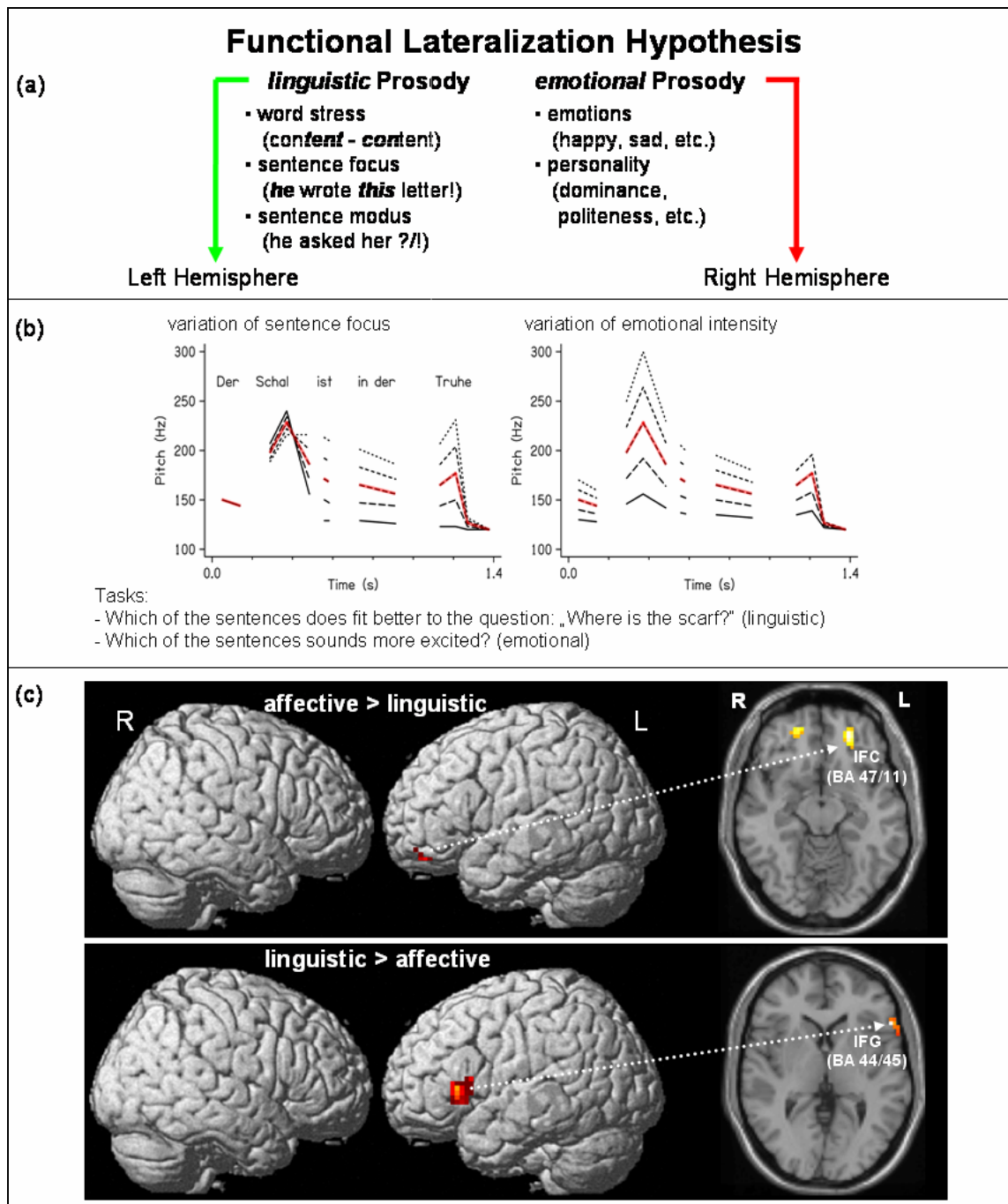


Fig. 4)

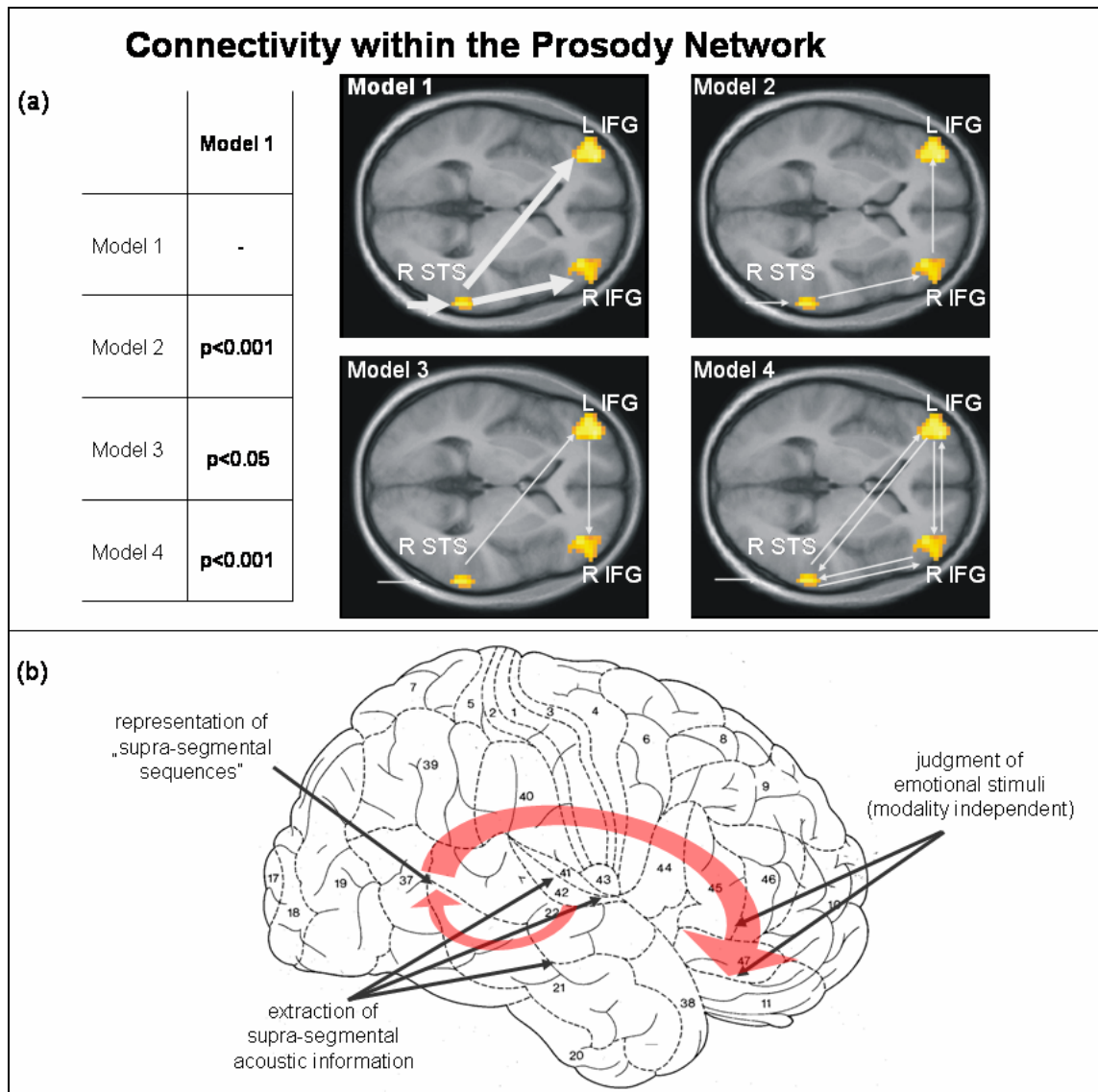


Fig. 5)

