# Prosody Production and Perception: Converging Evidence from fMRI Studies

*Jörg Mayer[1], Dirk Wildgruber[2], Axel Riecker[2], Grzegorz Dogil[1],*
*Hermann Ackermann[2], Wolfgang Grodd[2]*

[1]Experimental Phonetics, University of Stuttgart
[2]Neuroradiology and Neurology, University of Tübingen
Joerg.Mayer@IMS.Uni-Stuttgart.DE

## Abstract

Clinical observations of distortions of production and perception of prosody implicate that distinct, non-overlapping neural circuits are responsible for distinct prosodic cues and functions. These observations motivate a question whether similar evidence can be found in the neurologically intact brain. The experiment presented in this paper was constructed to check the neuroanatomical basis of the prosody generator, a functional unit in the phonological system which integrates and controls the variation of prosodic parameters [1]. The results show that relatively small, non-overlapping, fronto-basal areas of both the right and the left hemisphere are involved in the generation of prosody. Futhermore, we found clear evidence for the functional lateralization of prosody processing. Generating linguistically geared prosody revealed exclusively left hemisphere activation, while the production of affective prosody revealed right hemisphere activation only. Additionally, our results show interesting analogies with activation patterns of perception paradigms of recent functional imaging studies.

## 1. Introduction

Prosody is a mode of communication which provides a parallel channel to speech. Prosodic features, unlike other linguistic features, are often produced without conscious intention and are open to forms of interpretation which rely on emotional or non-cognitive processes. On the other hand, prosodic organization of human communication is continuous and highly correlated with the semantic, syntactic, morphological and segmental organization of speech. Regardless of function, there exist only three prosodically active phonetic parameters: duration, intensity and pitch.

The variety of prosody functions and cues in language processing has led to multiple hypotheses concerning the neurolinguistic and neuroanatomical basis of prosody. At least four hypotheses have been particularly influential (cf. [2] for a critical review).

(1) The right hemisphere hypothesis contends that all aspects of prosody are independently processed by the right hemisphere and integrated with the linguistic information (which is processed by the left hemisphere) via interhemispheric connections (i.e. the fibres of the corpus callosum) [3].

(2) The functional lateralization hypothesis assumes that there is a continuum from linguistic to affective functions of prosody and processing shifts from the left hemisphere (more linguistically-based tasks) to the right hemisphere (more affectively-based tasks) [4]. Recently presented evidence points out that linguistic prosody should possibly be further subdivided into phrase vs. syllable level prosody, with a left hemisphere dominance for short frames (syllable) and a right hemisphere dominance for long frames [5, 6].

(3) The subcortical processing hypothesis claims that prosodic functions are highly dependent on subcortical processing and are not lateralized to one or another hemisphere [7].

(4) The acoustic cues hypothesis contends that duration, pitch (and possibly intensity) may be independently lateralized [8].

All these contradictory hypotheses find their support mainly in the clinical observation of language and speech impaired subjects. There are problems with the interpretation of data from patients if it is used in isolation. This data may reflect neural reorganization or the development of compensatory strategies. It can not be simply assumed that the absence of function after a stroke means that the patient has normal cognition minus one part. Apart from that, the data provided by observation of patients are a product of a highly complex cognitive process which can be hardly further fractionalized. Modern cognitive theories question the assumption of a simple correspondence between complex tasks (like prosody) and large brain areas (like whole brain hemispheres).

In our research we assume a highly fractionalized and elaborated model of prosody generation (cf. Fig. 1) and test its individual components with experiments designed to reveal the function of the active, healthy brain. The main interest in the experiment described below concerns the role of prosodic function ("intonational meaning") and the influence of the size of address frames in prosody generation.

## 2. Methods

### 2.1. Subjects and Materials

The study recruited ten healthy native german subjects (five females, five males, mean age 26.2 years, range 21-32 years). All participants were right-handed as determined by standardized inventory, and none of them had a history of neurological disorders. Informed consent had been obtained from each subject. Subjects were paid for the participation in the experiment.

The experiment was designed in accordance with the cognitive subtraction paradigm. This paradigm requires a pair of tasks which are similar except for the presence ('activation') or absence ('control') of the one cognitive process that is examined. A general outline of the design ist given in Fig. 2.1. Under activation conditions, subjects were asked to produce a sentence-like sequence consisting of five syllables [dadadadada] with various pitch-accent types and locations (the FOCUS condition), with various boundary tone types (the MODUS condition), and with various kinds of emotional state marking (the AFFECT condition). As a control condition
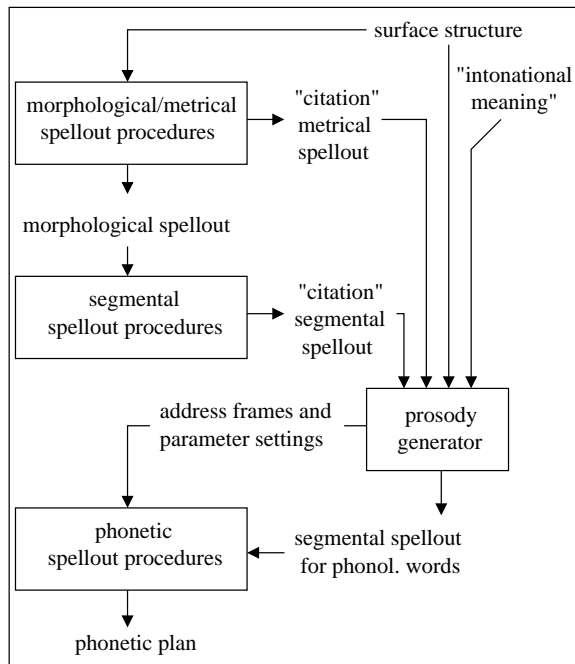
*Figure 1: The 'prosody generator' in Levelt's speech production model [1].*



*Figure 2: Cognitive decomposition of the experimental tasks (schematic representation).*

for the statistical analysis they were asked to produce the sequences [dadadadada, didididi, dodododo, dududududu] in a monotonous voice (with a syllable frequency of ca. 5 Hz; the MONO condition). Reactions were stimulated visually. Examples of the experimental material are given in Fig. 2.1.

We used reiterant syllables in order to reduce to the minimum the influence of the syntactic, semantic, morphological and segmental factors on prosody generation. The aspects of prosody that were controlled in this experiment were in accordance with the model of prosody generation given in [1], correlated only with different functions and different address frames and parameter settings (cf. Fig. 1). A further motivation for the use of reiterant speech was to avoid inhibitory activation of brain regions in the case of monotonous production of meaningful speech (control task).

### 2.2. Procedure

Subjects lie supine in the MR scanner (1.5 T whole body scanner, Siemens Vision), the heads being secured by means of a foam rubber in order to minimize movement artifacts. The stimuli were presented visually every 15 sec. for a period of three seconds. The pauses between the stimuli were 12 sec. long. Subjects were producing the required item immediate after stimulus presentation. Every 60 sec. there was a paradigm change, initiated by an acoustic instruction. Each stimulus has been presented eight times. In four out of these eight presentations the 'prosodic' reaction was required. In the other four cases the subjects were rendering the item in a monotonous manner. The material and the procedure have been validated in a pilot study performed outside of the MR scanner [9].

### 2.3. Imaging

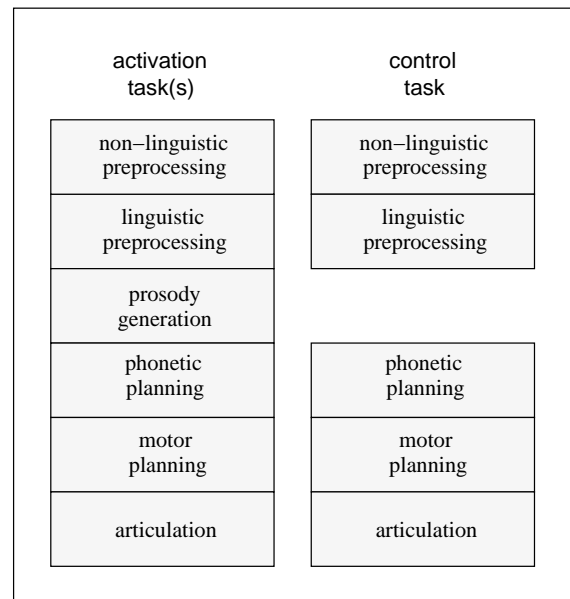fMRI [functional Magnetic Resonance Imaging] technology uses the blood oxygenation level dependent (BOLD) effect as an indirect marker of brain activation. Local neuronal activity gives rise to a decline in blood oxygenation which, in turn, causes an increase of blood flow. The hemodynamic response outweights the oxygen demand yielding accumulation of oxyhemoglobin within the respective region. Since magnetic properties of oxyhemoglobin are different from that of deoxyhemoglobin, imaging sequences allow to detect the change of MR signal within the activated areas. Twenty-eight parallel axial slices (thickness = 4 mm, gap = 1 mm) were acquired across complete brain volume by means of multislice echoplanar imaging sequence T2* EPI (TE = 39 ms, TR = 3 s, $\alpha = 90°$, FOV = 192 mm, $64^2$ matrix).

Imaging data was processed using SPM99 (Wellcome Institute of Cognitive Neuroscience, London). Functional images were movement corrected and coregistered with anatomical images. Spatial normalization procedures have been performed. The normalized fMRI data were spatially smoothed (Gaussian filter, 6 mm FWHM). The activation threshold was set at $p < 0.05$ (corrected).

## 3. Results

As expected, we observed a rather similar activation pattern for each condition compared to rest, given the virtually identical stimuli and task demands across conditions. For these contrasts, significant foci of activation were located in the motor cortex extending to premotor areas (MC), in the supplementary motor cortex (SMA), in the cerebellum (CER), and in occipital regions (OCC). These findings reflect the involvement of speech motor control (MC, SMA, CER) and processing of the visual stimuli (OCC).

In the direct contrasts between the prosodic conditions and their respective control condition these networks disappeared. This indicates that all tasks were comparable concerning the required cognitive components. As intended, the only difference between activation conditions and control conditions was the prosody processing part—active in the former and absent in the

Figure 3: Examples of stimulus/reaction pairs.



Figure 4: Projection of significantly activated voxels on the rendered mean anatomical images of all subjects.

latter. The subtraction FOCUS−MONO revealed a pattern of activation exclusively lateralized to the left hemisphere (Fig. 3). Foci of activation were located in the anterior part of the superior temporal gyrus (STG, BA 22) and in the inferior frontal gyrus (IFG, BA 47) (Fig. 3a). Additional clusters of activation were found in the left MC and in the anterior cingulum. The contrast MODUS−MONO revealed a focal spot of activation located also in the left IFG (BA 47) (Fig. 3 and Fig. 3b). Subtracting MONO from AFFECT we observed focal activation exclusively in the right IFG (BA 47/11) (Fig. 3 and Fig. 3c).

## 4. Discussion

Our results show that both cortical hemispheres subserve processing of prosody during speech production. This is clear evidence against a view that the right hemisphere is dominant for all types of prosodic information. Furthermore, our findings do not confirm a crucial role of subcortical structures in prosody processing since we did not observe any significant activation in that part of the brain. However, our results are consistent with the functional lateralization hypothesis. Generating linguistically controlled prosodic features on the syllable level (FOCUS) as well as on the phrase level (MODUS) revealed activation patterns exclusively in the left hemisphere, while generating different emotional expressions revealed exclusively right hemisphere activation.

Considering the exact localization of activated foci, our production data show interesting similarities with recent fMRI studies on prosody perception. Wildgruber et al. compared identification of prosodically expressed emotions (happiness, anger, fear, sadness, disgust) with a non-prosodic, phonological control task (identification of the vowel following the first "a" in a sentence) [10]. Subtraction of the control task from emotive identification elicited activation in the right STG and bilateral activation in the IFG (BA 47), with the right hemisphere showing overall greater activation in terms of magnitude and spatial extent. In a second experiment, Wildgruber et al. compared perception of affective prosody (rating of emotional expressiveness) with perception of linguistic prosodic features (rating of appropriateness of question-answer pairs with respect to accent placement; "semantic focus") [11]. Activation patterns of both
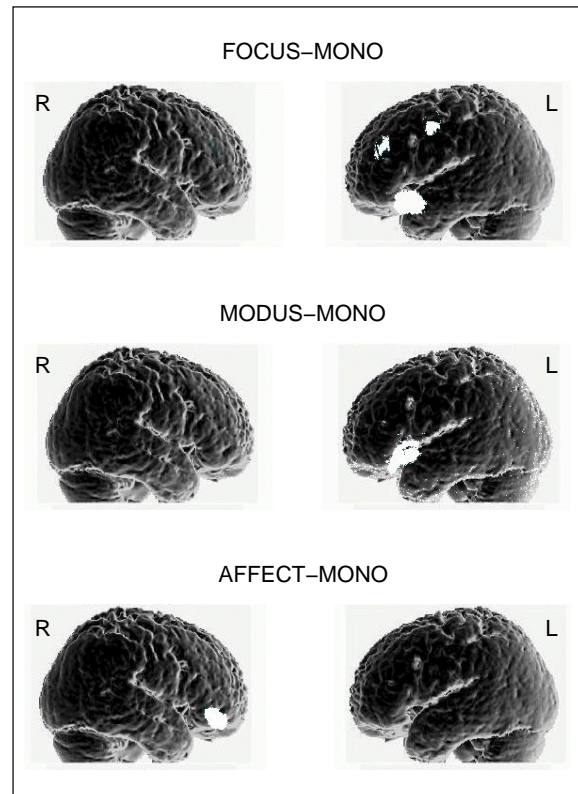
tasks involved again the IFG (BA 47/11). The linguistic task revealed additional activation in the left BA 44/45 ("Broca's area"). Focal activation in the orbito-basal frontal cortex during perceptual prosody processing was also shown by Meyer in a series of fMRI studies using filtered speech ('pure' intonation without segmental information) [6] and by Dapretto et al. using meaningful sentences in a selective attention paradigm [4].

Taken together, these findings point to a crucial role of the deep frontal operculum (BA 47 and 11) in the processing of the basic phonetic-phonological aspects of prosody. Besides adding evidence to the correlation between the function of prosody processing and the neuroanatomical region of the frontal operculum, our results show interesting similarities between speech production and perception: Both modi recruit the identical anatomical structures when certain functions are executed. Based on the referred studies and on our own data we can assume that the setting and interpretation of basic prosodic parameters is one such function. Only when these basic parameters have to be integrated with higher linguistic or paralinguistic information other areas become involved additionally. For example, integrating phonological and semantic focus activated Broca's area [11], while identification of emotions based on prosodic information activated the right STG (an analogous region to Wernicke's area in the left hemisphere) [10].

The role of the frontal operculum in prosody processing has to be further examined. Two directions seem particularly promising to us. First, detailed investigation of clinical populations with focal fronto-basal lesions should lead to a better understanding of the precise responsibility of this cortical area. Second, assuming that the frontal operculum is processing ba-

sic prosodic parameters, it should be tested, whether different acoustic cues give rise to differing activation patterns in that area.

# 5. References

[1] Levelt, W., 1989. *Speaking: From Intention to Articulation.* Cambridge, MA: MIT Press.

[2] Baum, S.; Pell, M., 1999. The neural bases of prosody. *Aphasiology*, 13(8), 581-608.

[3] Blumstein, S.; Cooper, W., 1974. Hemispheric processing of intonation contours. *Cortex*, 10, 147-158.

[4] Dapretto, M.; Hariri, A.; Bialik, M.; Bookheimer, S., 1999. Cortical correlates of affective vs. linguistic prosody: an fMRI study. *NeuroImage*, 9(6), S1054.

[5] Gandour, J.; Ponglorpisit, S.; Khunadorn, F., Dechongkit, S.; Boongird, P.; Boonklam, R.; Potisuk, S., 1992. Lexical tones in Thai after unilateral brain damage. *Brain and Language*, 43, 275-307.

[6] Meyer, M., 2000. *Ereignisbezogene hämodynamische Korrelate der Verarbeitung von grammatischer und prosodischer Information.* Doctoral thesis, University of Leipzig, Germany.

[7] Cancelliere, A.; Kertesz, A., 1990. Lesion localization in acquired deficits of emotional expression and comprehension. *Brain and Cognition*, 13, 133-147.

[8] Van Lancker, D.; Sidtis, J.J., 1992. The identification of affective-prosodic stimuli by left- and right-hemisphere-damaged subjects: All errors are not created equal. *Journal of Speech and Hearing Research* 35, 963-970.

[9] Mayer, J., 1999. Prosody processing in speech production: pre-evaluation of a fMRI study. *Proceedings 14th International Congress of Phonetic Sciences*, San Francisco, 2339-2342.

[10] Wildgruber, D.; Ackermann, H.; Klein, M.; Riecker, A.; Grodd, W., 2000. Brain activation during identification of affective speech melody: influence of emotional valence and sex. *NeuroImage*, 11(5), S341.

[11] Wildgruber, D.; Hertrich, I.; Ackermann, H.; Riecker, A.; Grodd, W., 2001. Processing of linguistic and affective aspects of speech intonation evaluated by fMRI. *NeuroImage*, 13(6), S627.
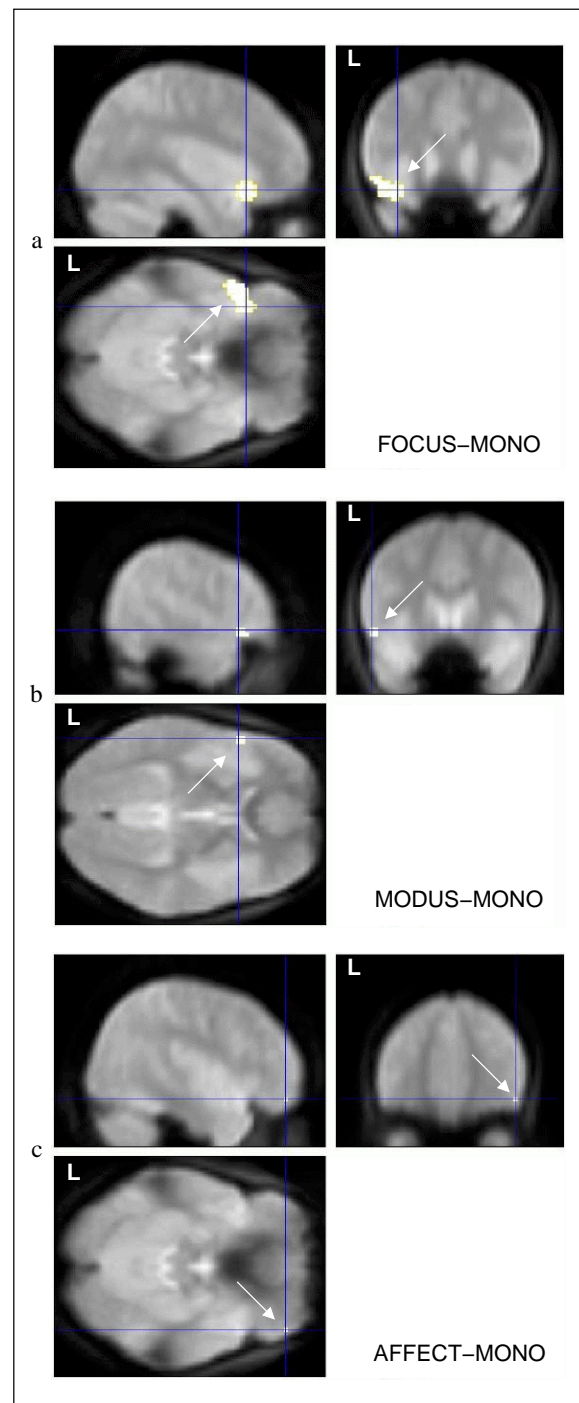
*Figure 5: SPMs of direct contrasts (prosodic conditions vs. control conditions); activated voxels are superimposed on a mean image of all functional images.*