Lateralized processing in human auditory cortex during the perception of emotional prosody

Beate Wendt[1,2], André Brechmann[1], Birgit Gaschler-Markefski[1], Henning Scheich[1] & Hermann Ackermann[2]

Special Laboratory Non-invasive Brain Imaging, Leibniz-Institute for Neurobiology[1]
Department of Neurology, University of Tübingen[2]
beate.wendt@ifn-magdeburg.de

Abstract

The aim of the present fMRI-study was to investigate the influence of different word prosodies on the activation of the auditory cortex (AC) of 24 subjects. Pseudowords and semantically neutral words were presented with neutral prosody in experiment I and with emotional prosodies in experiment II. We applied two lexical tasks i.e. detecting words or pseudowords. The control task was to detect pure tones. In both studies there was a typical left lateralized activation for speech perception on planum temporale (T3). This territory as part of Wernicke's area is specifically involved in speech perception.

A right lateralization simply dependent on prosodic versus neutral content of speech stimuli, as suggested by some literature, is not supported by the current results. In our experiments the emotional information was task-irrelevant and even distracted from the lexical task. Namely, the performance in the detection of words and pseudowords was significantly better in the prosodically neutral condition. Thus, the current results contribute to the clarification of the controversial issue whether prosodies lateralize brain activation to the right, i.e. if lexical rather than prosodic information is in the focus of a task involving prosodic material, a right hemisphere dominance cannot be expected.

1. Introduction

For the study of prosody perception different methods are used. Including dichotic listening tests, functional non-invasive brain imaging, and clinical studies e.g. on brain damaged patients [1; 11]. Based on these investigations, different theories regarding the laterality of prosody processing are discussed:

1. Acoustic hypothesis (bottom up)
   - Pitch differences are processed primarily in the right hemisphere, durational parameters, on the left side [19; 27; 30].
   - The fundamental frequency is processed in the right hemisphere and integrated with semantic and syntactic information from the left hemisphere via the corpus callosum. It is assumed that the acoustic parameter duration is processed in the left hemisphere. [17]
   - Deficits in the processing of spectral and temporal acoustic information are responsible for the disturbed perception of prosody [8]
2. Functional hypothesis (top down)
   - Emotional aspects of prosody are processed in the right hemisphere whereas linguistic aspects of prosody are processed on the left side [4]
   - Prosody processing takes place in right hemisphere regions which on the left side are responsible for language processing. [22; 17; 26; 23]

Non-uniform results were found regarding the functional role of the hemispheres in prosody processing. One important reason for this may be the use of language material. In most studies on prosody, semantic and syntactic processing involves a stronger activation of the left hemisphere. This additional activation of the left hemisphere could possibly conceal the processing of prosodic features in the right hemisphere [1]. A further reason for these inconsistencies may be methodological differences between the studies. Often, the prosodic stimulus material is not sufficiently validated [19], so that the intended emotion can not be identified even by healthy control subjects [25].

For the reasons stated above, it is presently unclear whether the processing of prosody is a dominant function of one hemisphere or only a gradual difference in the effectiveness of the processing between the hemispheres.

The aim of the present study was to test the lateralization of activation in auditory cortex in a lexical decision task involving speech stimuli with either neutral or emotional prosody taken from a prosody corpus developed and evaluated in our lab.

2. Corpus

In a communicational situation the interpretation of prosodic expressions are influenced by factors like gestures, facial expression, and conversational context. In studies on auditory processing of prosody perception all these aspects of a communicational situation lead to additional, confounding brain activation. Therefore, it is not possible to disentangle the effect produced by the auditory processing of prosody. Hence, stimuli for such studies must be suitable for prosody and speech perception tests without communication context.

Therefore, we developed the "Corpus of spoken words for studies of auditory speech and emotional prosody processing" (WaSeP©) for the investigation of speech and prosody perception under the special requirements of non-invasive brain imaging [29].

2.1. First Part - words

The first part of the corpus contains more than 3000 two-syllabic German nouns with one master morpheme in nominative singular form, spoken by an actor and an actress in standard German pronunciation. All stimuli were recorded in an acoustically controlled environment. All nouns of WaSeP© were evaluated for there lexical emotional connotation by 36 German native speakers (from 19 to 52 years). For each noun, the number of allocations to the different semantic categories (positive, neutral, negative) was registered and a rank scaling was generated. Those nouns, which occupied the first one hundred ranks of either the positive, neutral, or negative scale, were encoded [encoding - 2] to convey the target emotions joy, sadness, anger, fear, and disgust.

2.2. Second Part - pseudowords

The second part of corpus consists of 222 two-syllabic pseudowords, which resulted from permutations of the two-syllabic German nouns from the first part. These pseudowords, which were phonetically balanced and corresponded to the phonological, phonotactic and phonetic rules of German language, were spoken by the actors with the following prosodies: neutral, joy, sadness, anger, fear, and disgust. The emotional prosodic expressions of these stimuli were evaluated by a phonetically untrained group of 74 native German listeners (age 18 to 62 years) [decoding – 2].

All emotional prosodic expressions were identified by more than 70% of all listeners. There were no significant differences with respect to the gender of the listeners or the speakers.

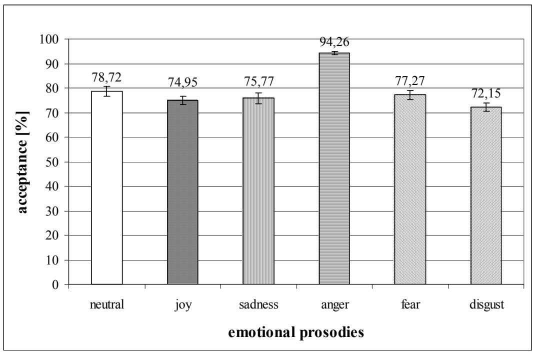

Figure 1: *Decoding emotional prosodies of the untrained group (74).*

3. Method

3.1. Subjects

Twenty four healthy, normal hearing native German speakers (12 female and 12 male, age 21 – 40), participated in the experiments. They gave written informed consent to the study which was approved by the ethics committee of the University of Magdeburg.

3.2. Data acquisition

Before the fMRI session, participants answered the hearing questionnaire from the WaSeP© and thus were familiar with the stimuli. Furthermore, they were experienced in the scanning procedure due to previous participation in several other fMRI studies.

Low noise fMRI experiments were carried out in a BRUKER 3T/60 head scanner equipped with a quadrupolar birdcage head coil [24]. Pilot scans were used for orientation of 4 contiguous 8mm slices covering the superior temporal plane in both hemispheres by following the course of the sylvian fissure on both sides as closely as possible.

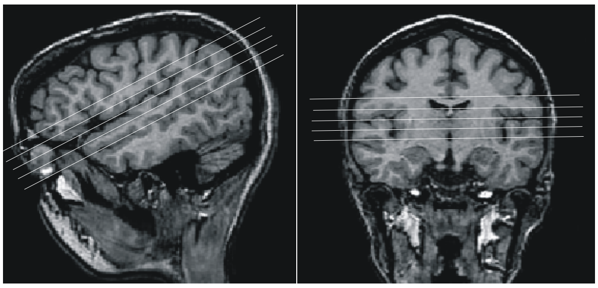

Figure 2: *Slice orientation at fMRI experiments.*

During functional imaging, pseudowords and semantically neutral words from the WaSeP© were presented binaurally with neutral prosody in experiment I and with emotional prosodies in experiment II via electro-dynamic headphones [3]. The control task was to detect pure tones (440 Hz). Each stimulus block of 24 sec consisted of words, pseudowords and pure tones in randomized order and alternated with silence blocks of 24 sec.

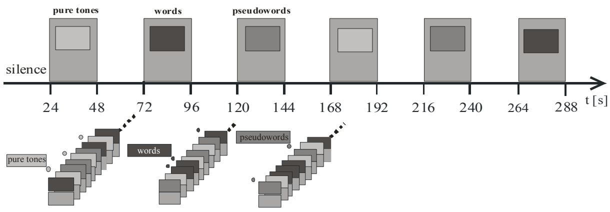

Figure 3: *Block design of both experiments.*

For each experiment, 215 functional images of each slice were acquired in 17 min12 sec using a FLASH sequence (TE: 32 ms; TR: 179.25 ms; flip angel: 15°; matrix size: 64x46; field of view: 18 cm). We used the keyhole technique to increase data acquisition rate (keyhole-factor: 0.5, keyhole-block-size: 5).

3.3. Data processing

Each functional data-set was subjected to a quality check. The data processing was monitored using the AIR package. Images were corrected for 2D-movements using the AIR package. Functional data were analyzed with the software package KHORFu [13]. Activated voxels ($p < 0.001$) were assigned to the four territories TA, T1, T2, and T3 which were defined in each individual subject by using a combination of anatomical landmarks and clusters of fMRI activation [6]. The intensity weighted volume (IWV), as the product of the number of activated voxels and their mean BOLD signal intensity change, was computed.

For stimulus presentation and recording of behavioral responses the software Presentation (Neurobehavioral Systems, Inc.) was used. During the fMRI experiments, the reaction time (key pressing), number of correct responses, false alarms, misses, and correct rejections were recorded. From these data, the sensitivity index - d’ was computed for each subject.

4. Results

4.1. Experiment I

The sensitivity index revealed significant differences in performance in the detection of words and pseudowords ($p<0.001$) as well as words and pure-tones ($p<0.05$). Subjects detected words significantly better than pseudowords and pure tones.

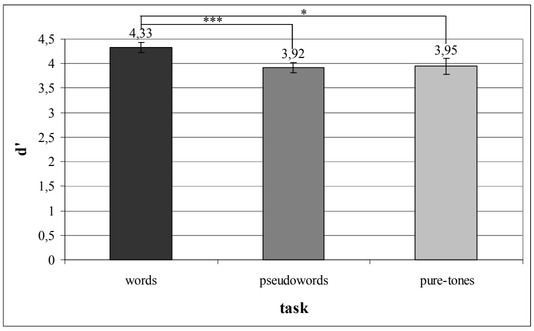

Figure 5: *Sensitivity index in study I.*

All three conditions led to bilateral auditory cortex (AC) activation with respect to silence blocks significantly stronger activation in left than in right AC.

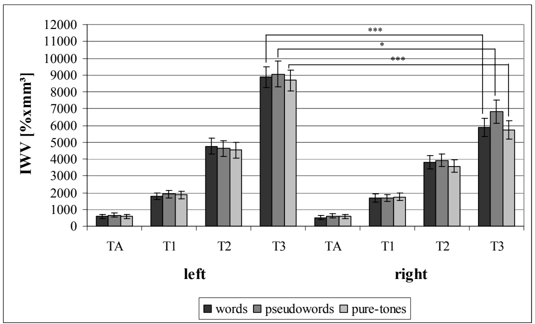

Figure 6: *Activation in auditory territories in study I.*

This effect was mainly due to a significantly left lateralized activation of territory T3 on planum temporale (words: $p<0.001$, pseudowords: $p<0.05$, pure tones: $p<0.001$).

4.2. Experiment II

The sensitivity index revealed significant differences in performance in the detection of words and pseudowords ($p<0.001$) as well as pseudowords and pure-tones ($p<0.001$). Subjects detected pseudowords significant worse than words and pure-tones.

The fMRI activation in auditory cortex was similar to that of experiment I. Again, all three conditions led to significantly stronger activation in left than in right AC ($p<0.05$) which was mainly due to left lateralized activation in T3 (words: $p<0.001$, pseudowords: $p<0.01$, pure tones: $p<0.05$).

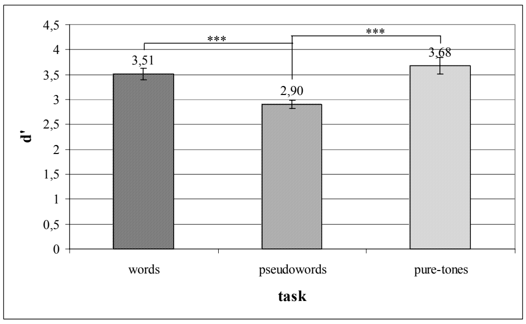

Figure 7: *Sensitivity index in study II.*

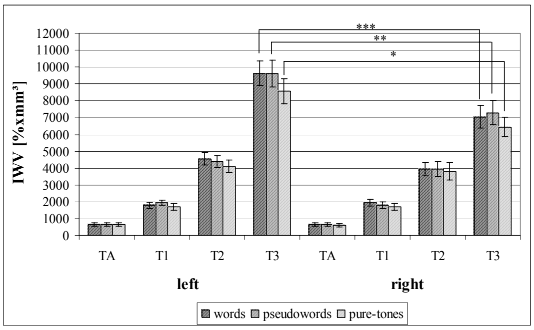

Figure 8: *Activation in auditory territories in study II.*

4.3. Comparison between Experiment I and II

The sensitivity index revealed significant differences in performance of experiment I and II. The detection of words as well as pseudowords in experiment II was significant worse ($p < 0.001$). The detection of pure tones, however, was not significantly different between the experiments.

The differences in performance in the lexical decision tasks are not reflected in the activation of the auditory cortex (AC). There were no significant differences between experiment I and II, neither for the global activation nor for any individual territory.

5. Conclusions / Discussion

In both experiments, all three conditions led to left-lateralized activation in auditory cortex (AC). This was mainly due to activation in auditory territory T3 on planum temporale which is part of the Wernicke's area. This sensory speech area has recently been described as being involved in further processing of perceived auditory stimuli and integration into available knowledge [5; 10].

Activation of left and right AC was also independent of the stimuli being prosodically modulated or neutral. On the one hand it is possible, that prosodic information processing is not strongly lateralized to the right hemisphere as suggested by some authors [11; 18; 20] describing a bilateral network for the processing of prosodic information. On the other hand the stimulus independent results may be due to the fact that solving the tasks did not rely on the emotional prosodies but only on the lexical information of the tones. This finding is consistent with results from our lab showing that fMRI activation is strongly task dependent [7]. There it was shown that presenting exactly the same set of stimuli with two different tasks can lead to a shift in laterality of auditory cortex activation. Categorization of the direction of frequency-modulated tones (rising vs. falling) strongly involved the right auditory cortex whereas categorization of the duration of the same stimuli mainly involved the left auditory cortex. Thus, it has to be tested in further experiments whether a task involving the identification of specific emotional prosodies of the same speech stimuli used in the present study would lead to a stronger activation of the right auditory cortex.

6. Acknowledgments

This study was supported by the Deutsche Forschungsgemeinschaft (SFB 426).

7. References

[1] Ackermann, H., I. Hertrich, et al. (1993). "Prosodische Störungen bei neurologischen Erkrankungen - eine Literaturübersicht." Fortschr. Neurol. Pschiat. 61: 241-253.

[2] Banse, R. und K. R. Scherer (1996). "Acoustic profiles in vocal emotion expression." J Pers Soc Psychol 70(3): 614-36.Baumgart et al. 1998; Med. Phys. 25(10): 2068-2070

[3] Baumgart, F., T. T. Kaulisch, C., et al. (1998). "Electrodynamic headphones and woofers for application in magnetic resonance imaging scanners." Med. Phys. 25(10): 2068-2070.

[4] Behrens, S. J. (1985). "The perception of stress and lateralization of prosody." Brain Lang 26(2): 332-48.Brechmann, A. et al (2002), J Neurophysiol 87: 4230-4233

[5] Binder, J. (2000). "The new neuroanatomy of speech perception." Brain 123(Pt 12): 2371-2

[6] Brechmann, A., F. Baumgart, et al. (2002). "Sound-level-dependent representation of frequency modulations in human auditory cortex: a low-noise fMRI study." J Neurophysiol 87(1): 423-33.

[7] Brechmann, A. and H. Scheich (2005). "Hemispheric shifts of sound representation in auditory cortex with conceptual listening." Cereb Cortex 15(5): 578-87.

[8] Cooper, W. E. (1981). "The analytic/holistic distinction applied to the speech of patients with hemispheric brain damage." The Behavioral and Brain Sciences 4: 68-69.

[9] Friederici, A. D. (2000). "The developmental cognitive neuroscience of language: a new research domain." Brain Lang 71(1): 65-8.

[10] Friederici, A. D. und D. Y. von Cramon (1999). Neurobiologische Grundlagen des Sprachverstehens. Sprachrezeption. A. D. Friederici. Göttingen, Bern, Toronto, Seattle, Hogräve-Verlag für Psychologie. Band 2: 307-349.

[11] Gandour, J., Y. Tong, et al. (2004). "Hemispheric roles in the perception of speech prosody." Neuroimage 23(1): 344-57.

[12] Gao, J. H., J. Xiong, et al. (1996). "Improving the temporal resolution of functional MR imaging using keyhole techniques." Magn Reson Med 35(6): 854-60.

[13] Gaschler, B., F. Schindler, et al. (1996). KHORFu: a KHOROS-based functional image post processing system. A statistical software package for functional magnetic resonance imaging and other neuroimage data sets. COMPSTAT 1996: proceedings in computational statistics. 12th symposium held in Barcelona, Spain, 1996. A. Prat. Heidelberg, Physica-Verlag: 57-58.

[14] Hartje, W. (2002). Funktionelle Asymmetrie der Großhirnhemisphären. Klinische Neuropsychologie. W. Hartje und K. Poeck. Stuttgart; New York, Thieme Verlag: 67-92.

[15] Heilman, K. M., R. Scholes, et al. (1975). "Auditory affective agnosia. Disturbed comprehension of affective speech." J Neurol Neurosurg Psychiatry 38(1): 69-72.

[16] Jones, R. A., O. Haraldseth, et al. (1993). "K-space substitution: a novel dynamic imaging technique." Magn Reson Med 29(6): 830-4.

[17] Klouda, G. V., D. A. Robin, et al. (1988). "The role of callosal connections in speech prosody." Brain Lang 35(1): 154-71.

[18] Kotz, S. A., M. Meyer, et al. (2003). "On the lateralization of emotional prosody: An event-related functional MR investigation." Brain Lang 86(3): 366-76.

[19] Myers, P. (1999). Prosodic Deficits. Right Hemisphere Damage - Disorders of Communication and Cognition. San Diego, London, Singular Pablishing Group, Inc.: 73-90.

[20] Pell, M. D. (2005). "Cerebral mechanisms for understanding emotional prosody in speech." Brain Lang.

[21] Poeppel, D. (2003). "The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'." Speech Communication 41: 245-255.

[22] Ross, E. D. und M. M. Mesulam (1979). "Dominant language functions of the right hemisphere? Prosody and emotional gesturing." Archives of Neurology 36: 144-148.

[23] Ross, E. D., R. D. Thompson, et al. (1997). "Lateralization of affective prosody in brain and the callosal integration of hemispheric language functions." Brain and Language 56: 27-54.

[24] Scheich, H., F. Baumgart, et al. (1998). "Functional magnetic resonance imaging of a human auditory cortex area involved in foreground-background decomposition." Eur. J. Neurosci. 10: 803-809.

[25] Tompkins, C. A. (1991). "Automatic and effortful processing of emotional intonation after right or left hemisphere brain damage." J Speech Hear Res 34(4): 820-30.

[26] Tucker, D. M., R. T. Watson, et al. (1977). "Discrimination and evocation of affectively intoned speech in patients with right parietal disease." Neurology 27(10): 947-50.

[27] Van Lancker, D. and J. J. Sidtis (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.

[28] van Vaals, J. J., M. E. Brummer, et al. (1993). ""Keyhole" method for accelerating imaging of contrast agent uptake." J Magn Reson Imaging 3(4): 671-5.

[29] Wendt, B. and H. Scheich (2002). The "Magdeburger Prosodie Korpus" - a spoken language corpus for fMRI-Studies. Speech Prosody 2002, Aix-en-Provence, SProSIG. [http://www.ifn-magdeburg.de/en/special_labs/non-invasive_brain_imaging/projects/2/4/index.jsp]

[30] Zatorre, R. J., P. Belin, et al. (2002). "Structure and function of auditory cortex: music and speech." Trends Cogn Sci 6(1): 37-46.