

Brain activity varies with modulation of dynamic pitch variance in sentence melody

Martin Meyer,^{a,b,*} Karsten Steinhauer,^{c,d} Kai Alter,^a Angela D. Friederici,^a and D. Yves von Cramon^a

^a *Max-Planck-Institute of Cognitive Neuroscience, Leipzig, Germany*

^b *Department of Neuropsychology, University of Zürich, Treichlerstrasse 10, CH-8032 Zürich, Switzerland*

^c *Brain and Language Lab, Georgetown University, Washington DC, USA*

^d *School of Communication Sciences and Disorders, McGill University, Montreal, Canada*

Accepted 20 August 2003

Abstract

Fourteen native speakers of German heard normal sentences, sentences which were either lacking dynamic pitch variation (flattened speech), or comprised of intonation contour exclusively (degraded speech). Participants were to listen carefully to the sentences and to perform a rehearsal task. Passive listening to flattened speech compared to normal speech produced strong brain responses in right cortical areas, particularly in the posterior superior temporal gyrus (pSTG). Passive listening to degraded speech compared to either normal or flattened speech particularly involved fronto-opercular and subcortical (Putamen, Caudate Nucleus) regions bilaterally. Additionally the Rolandic operculum (premotor cortex) in the right hemisphere subserved processing of neat sentence intonation. As a function of explicit rehearsing sentence intonation we found several activation foci in the left inferior frontal gyrus (Broca's area), the left inferior precentral sulcus, and the left Rolandic fissure. The data allow several suggestions: First, both flattened and degraded speech evoked differential brain responses in the pSTG, particularly in the planum temporale (PT) bilaterally indicating that this region mediates integration of slowly and rapidly changing acoustic cues during comprehension of spoken language. Second, the bilateral circuit active whilst participants receive degraded speech reflects general effort allocation. Third, the differential finding for passive perception and explicit rehearsal of intonation contour suggests a right fronto-lateral network for processing and a left fronto-lateral network for producing prosodic information. Finally, it appears that brain areas which subserve speech (frontal operculum) and premotor functions (Rolandic operculum) coincidentally support the processing of intonation contour in spoken sentence comprehension.

© 2003 Elsevier Inc. All rights reserved.

Keywords: Functional MRI; Dynamic pitch variation; Sentence prosody; Peri-sylvian cortex; Planum temporale; Frontal operculum; Rolandic operculum; Basal ganglia; Language and motor integration; Auditory rehearsal

1. Introduction

Comprehending spoken language includes the decoding of information from differing linguistic domains, e.g., semantics of words, thematic and structural relations, as well as from nonlinguistic and linguistic acoustical cues, commonly referred to as 'prosody.' Prosody describes abstract phonological phenomena such as word stress, sentence accent, and phrasing and

refers also to the phonetic attributes used to encode these abstract structures, i.e., intonation, amplitude, duration, etc. Listeners can elicit information from intonation, duration, and amplitude to help decode the syntactic and focus structure of the sentences they attend to (Steinhauer, 2003; Steinhauer, Alter, & Friederici, 1999). Thus, prosody has a linguistic function at many different levels. During speech comprehension it contributes to the interpretation of the linguistic signal. Modulation of prosodic parameters, i.e., of pitch accent, guides syntactic parsing even though pitch accent per se is not a syntactic property. Slow pitch movements which

* Corresponding author. Fax: +41-1-634-4342.

E-mail address: m.meyer@psychologie.unizh.ch (M. Meyer).

extend over chunks of utterances longer than just one segment constitute the intonation contour, i.e., speech melody. In intonational languages such as German, French, and English, speakers use different pitch movements to signal questions vs. statements, where questions are usually produced with a pitch rise at the end of the utterance.

Recent brain imaging studies identified distinct brain regions, especially in left peri-sylvian cortex subserving particular aspects of sentence-level speech comprehension (Bookheimer, 2002; Friederici, 2002; Friederici & Alter, this volume; Friederici, Rüschemeyer, Hahne, & Fiebach, 2003b; Kaan & Swaab, 2002). Neither the spectral and temporal features that carry the relevant attributes of prosody, nor the cerebral substrates of prosodic parameters (e.g., speech melody) available in spoken language, however, have been exactly identified so far (Lakshminarayanan et al., 2003). Patients suffering from either left or right hemispheric lesions showed comprehension deficits for linguistic intonation giving credence to the view that prosodic processing is mediated by both hemispheres (Pell & Baum, 1997). This view lends support to a recent model which proposed that prosodic functions are not localized in either the right or left hemisphere (Dogil et al., 2002). According to this model the prosodic frame length rather than prosody per se dictates the lateralization. Prosodic features which require a short address frame, (e.g., focused syllable) are lateralized differently as compared to prosodic elements comprising a long address frame, (e.g., intonational phrases). In complement to this, the ‘asymmetric sampling in time hypothesis (AST)’ argues in favour of a functional hemispheric difference which derives from the manner in which auditory signals are processed at an early stage (Pöppel, 2003). This hypothesis holds that speech processing even at an early stage generally occurs symmetrically in both the left and the right hemisphere. However, the signal is elaborated differentially in the time domain. Left non-primary auditory areas extract information from short (20–50 ms), and right hemisphere homologues pick up information from long temporal integration windows (150–250 ms). Linguistically, the ‘AST’ hypothesis suggests that prosodic processing at the level of lexical stress is lateralized to the left hemisphere. Additionally, right temporal regions are more proficient at processing prosody at the level of intonation contour. Empirical evidence supporting this view comes from a recent lesion study on patients who underwent a resection of right superior temporal areas and unveiled an impairment at using pitch contour information (Johnsrude, Penhune, & Zatorre, 2000). Additionally, an fMRI-study from our lab demonstrated a stronger right hemisphere involvement, particularly in the right planum temporale (PT) and the right Rolandic operculum (ROP) in processing slow prosodic modulations, (e.g., pure sentence intonation)

(Meyer, Alter, Friederici, Lohmann, & von Cramon, 2002). In this study, volunteers heard either normal sentences or pure sentence intonation degraded speech. The latter condition was derived from normal sentences which underwent a special filtering procedure (PURR-Filter (cf. Section 2) which removes segmental, but preserves prosodic parameters, (i.e., intonation, duration, and amplitude envelope) which represent speech melody. However, in this study we contrasted normal speech to degraded speech both comprising intonation contour. In order to elucidate the cerebral regions subserving prosodic variation in spoken language we designed a new fMRI-study.

1.1. The present study

Here, participants were presented with three conditions, *normal sentences*, *degraded speech*, and *flattened speech*. The latter was derived from normal sentences by selectively removing variations in the intonation contour (fundamental frequency, or F_0), but preserving other prosodic parameters, (segmental/phonemic information, amplitude envelope, and duration). This design allowed us to contrast speech signals comprising intonation (*normal speech*, *degraded speech*) to speech signals without any sentence intonation (*flattened speech*). Based on the results of our previous study we might expect larger activation for normal sentences and flattened speech in the peri-sylvian cortex, namely the left inferior frontal gyrus (IFG) and the superior temporal gyrus (STG) bilaterally when compared to degraded speech, since the latter lacks any lexical and syntactic information. We predicted stronger brain responses to degraded speech in the fronto-opercular areas including the basal ganglia bilaterally reflecting an interaction between speech processing and related effort (Meyer et al., 2002). Additionally we expected activation in the right inferior premotor cortex for degraded, but not for normal speech pointing to a role this particular area plays in implicit sub-articulation whilst participants heard speech melody. As phonological working memory has been found to involve frontal structures particularly of the left hemisphere (Paulesu, Frith, & Frackowiak, 1993), the right hemisphere activation may shift to left homologues when volunteers perform an explicit rehearsal task. We therefore asked the subjects to explicitly hold in memory intonation contour. Separate analyses of the time windows for passive listening and for active rehearsal of the sentence contour should enable to distinguish between the networks for perceiving and reproducing sentential melodies. Finally, as far as the two models mentioned above are concerned we intended to test whether the planum temporale (PT) can be identified as the brain region that mediates the integration of slowly as well as rapidly changing acoustic cues during comprehension of spoken language. Even

though the functional role the PT plays has not yet been specified there is some evidence that it mediates processes of phonetic integration. Generally, the PT is assumed to constitute any complex auditory processing, including speech, at an early stage including speech (Binder et al., 2000; Hickok & Pöppel, 2000). In a recent review paper it has been suggested that the planum temporale acts as a *computational hub* for the integration of spectrotemporal patterns (Griffiths & Warren, 2002). Jäncke, Wüstenberg, Scheich, & Heinze (2002) recently provided compelling evidence for the view that the PT is not particularly involved in phonetic analysis per se, but is rather specialized for the analysis of rapidly changing cues. Here, we would like to augment this view by elucidating the role of the PT regarding slowly as opposed to rapidly changing acoustic cues in spoken language. If the PT is crucial for integration of slow changes (i.e., F_0 variations), we may expect larger PT activity for prosodically adverse stimuli *degraded speech*, *flattened speech* when compared to normal sentences, reflecting the violation of prosodic integrity.

2. Materials and methods

2.1. Subjects

Fourteen native German volunteers (6 male, age range 18–27, mean 22.7) participated in the study after giving written informed consent in accordance with the guidelines approved by the Ethics Committee of the Leipzig University Medical Faculty. Volunteers were assessed as right-handed according to the *Edinburgh Handedness Inventory* (Oldfield, 1971). Participants had no hearing or neurological disorders and normal structural MRI scans. They had no prior experience with the task and were not familiar with the stimulus material.

2.2. Stimuli

The German sentence material consisted of 144 stimuli (72 natural sentences and 72 artificially manipulated sentences) varying pitch parameters. All speech signals were controlled for their duration and loudness.

1. *Normal speech*. This condition includes three subtypes of sentences which differ slightly in their intonational contour.¹ All sentences were infinitival sentences containing a control verb such as ‘promises’ and an infinitive (see example below) varying the position of sentence accents which either appeared on the first noun phrase, on the second noun phrase or on the first verb. Since the distinct sub-conditions are not assumed to yield substantial hemodynamic differences, all sub-con-

ditions were included for reasons of a highly desirable controlled variability of natural speech input.

PETER verspricht *ANNA* zu *ARBEITEN* und das Büro zu putzen.

PETER promises *ANNA* to *WORK* and to clean the office.

2. *Flattened pitch*. The flattened pitch condition was derived by using a special speech re-synthesis procedure² in order to generate a violation of the sentence prosody. All normal sentences were manipulated by re-synthesis. The manipulation is based on an algorithm (WinPitch, cit.) allowing the re-adjustment of the F_0 -contour. The speech re-synthesis was carried out at the mean F_0 -value of the speaker’s voice, i.e., at 200 Hz by applying a simple linear function between onset and offset of each sentence. This procedure removes the original geometrical characteristics such as linguistically triggered pitch accents and the declination line. In addition, global slow modulations had been removed yielding a monotonous sounding sentence. These modulations concerned the pitch contour varying over domains which size is larger than one syllable. Apart from the pitch contour (F_0) the re-synthesis procedure preserves both syllabic and rapidly changing sub-syllabic varying information (e.g., amplitude envelope, duration) in the speech signal. Fig. 1 illustrates that the resulting signal does not contain any dynamic pitch variations, i.e., no peaks and valleys, and is thus flattened globally. On the other hand the manipulated speech signal contains all information necessary to perform phoneme detection, lexical access, syntactic and semantic processing.

3. *Degraded speech*. To achieve a speech signal which lacks lexical and syntactic information normal sentences were subjected to PURR-filtering procedure (Sonntag & Portele, 1998). PURR-manipulated speech stimuli contain only prosodic parameters such as intonation, duration, amplitude envelope, and the second and third harmonics. From a linguistic point of view, these phonetic attributes represent speech melody, the distribution and type of pitch accents and boundary markers of prosodic domains. In more perceptual terms *degraded speech* could be described to sound like speech melody listened to from behind a door. Fig. 1 demonstrates that acoustic information exceeding the 3rd harmonic had been removed from the speech signal.

For technical reasons it was not possible to add a fourth condition [-prosody] and [-segmental] which is supposed to complete the design as the PURR filter was disrupting the re-synthesized speech files. All normal sentences were recorded with a trained female speaker in a sound proof room (IAC) at a 16 bit/41.1 kHz sampling rate and then digitised. Both flattened and degraded speech were derived from normal sentences. Since the PURR-procedure allows for only a 16 bit/

¹ Example sound files are available at “http://www.psychologie.unizh.ch/neuropsy/home_mmeyer/YBLRN2956”.

² WinPitch 1.89, Pitch Instruments, Toronto, Ont., Canada.

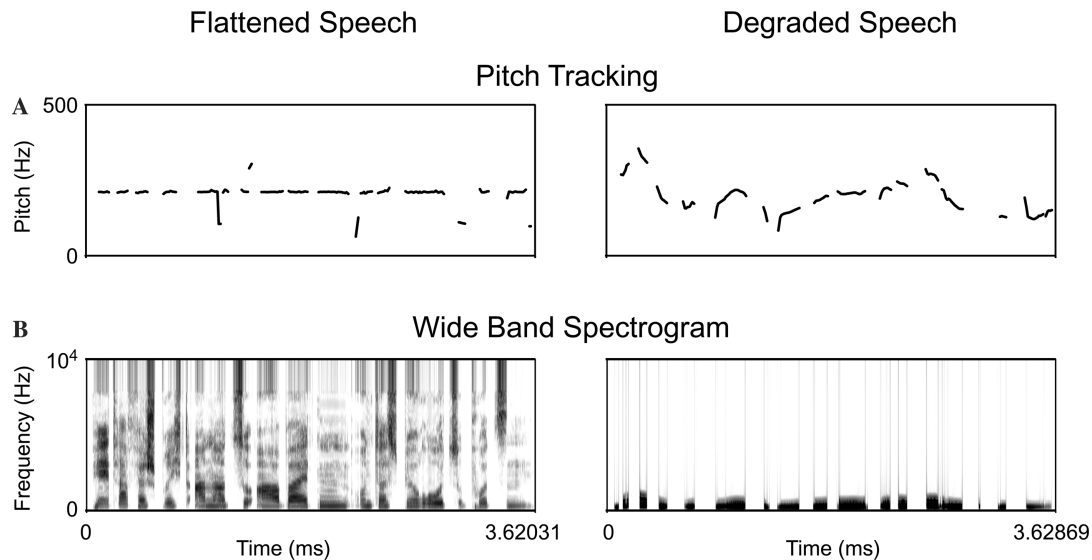


Fig. 1. Acoustic analyses of flattened and degraded speech. Speech signals of sentence derived from resynthesis revealed flattened pitch contour (A), but unaltered wide band spectrum of frequencies (0–10 kHz) (B). The artificial re-synthesis eliminates prosodic cues in an intonational language such as German, i.e., the typical rising and falling F_0 pattern over the whole sentence. The right side of the figure shows normal sentence intonation for degraded speech (A) whilst the wide band spectrogram illustrates reduced frequency information of a degraded sentence (B).

16 kHz sampling rate, all signals to be presented in the experiment were downsampled to avoid quality differences across conditions. All stimuli except degraded signals were normalized in amplitude (70%). Since the latter were limited in bandwidth compared to the other three conditions a stronger normalization (85%) was necessary to guarantee equal loudness. The mean length of the sentences in the ‘natural speech’ condition was $\pm SD$, 3.61 ± 0.23 s, in the ‘flattened pitch’ condition $\pm SD$, 3.85 ± 0.28 s, and in the delexicalized ‘degraded speech’ condition $\pm SD$, 3.81 ± 0.28 s.

2.3. Procedure

Participants heard the stimuli occurring in pseudo-random order.³ Sentences were not repeated during the experiment. The sounds were presented binaurally via specially constructed headphones. The study employed a single-trial design to enable an event-related analysis (D’Esposito, Zarahn, & Aguirre, 1999). To allow the hemodynamic response to return to baseline level adequately each sentence was followed by an Inter-Trial-Interval lasting twelve seconds until the onset of the following trial. The entire experimental session consisted of two blocks (runs), each comprising 72 trials.

2.4. Task

Participants were asked to perform in a prosody comparison task. First, they had to closely listen to the

sentence intonation and to rehearse this percept during the Inter-Stimulus Interval. Whenever a trial was unpredictably initially marked as a ‘Compare’ trial, subjects had to judge whether the current and the preceding sentence shared the same prosodic pattern (*yes/no* judgement). The number of ‘Compare’ trials was reduced to 20% (i.e., 7 trials in each condition) in order to avoid a general influence of this matching task on sentence perception and not to overtax working memory. ‘Compare’ trials were included in the design only to guarantee that language stimuli were attended to appropriately. However, ‘Compare’ trials were excluded from further analysis in order to avoid confounding between rehearsal and template matching processes. At the beginning of the remaining 80% of trials (28 trials in each condition), the presentation of a sine wave tone (1000 Hz, 100 ms) indicates that no explicit comparison was required.

2.5. MRI data acquisition

MRI data were collected at 3T using a Bruker 30/100 Medspec system (Bruker Medizintechnik GmbH, Ettlingen, Germany). For each subject, structural and functional (echo-planar) images were obtained from eight axial slices parallel to the plane intersecting the anterior and posterior commissures (AC–PC plane). The most inferior slice was positioned below the AC–PC plane and the remaining seven slices extended dorsally. The whole range of slices comprised an anatomical volume of 46 mm and covered all parts of the peri-sylvian cortex and extended dorsally to the intraparietal sulcus. After defining the slices’ position a set of two-dimensional T_1 weighted anatomical images (IR-RARE

³ A comprehensive report on methods and procedure is available at “http://www.psychologie.unizh.ch/neuropsych/home_mmeyer/YBLRN2956”.

sequence: TE = 20 ms, TR = 3750 ms, in-plane resolution 0.325 mm²) were collected in plane with the echo-planar images, to align the functional images to the 3D-images. A gradient-echo EPI sequence was used with a TE 30 ms, flip angle 90°, TR 2 s, acquisition bandwidth 100 kHz. The matrix acquired was 64 × 64 with a FOV of 19.2 cm, resulting in an in-plane resolution of 3 mm × 3 mm. The slice thickness was 4 mm with an interslice gap of 2 mm. In a separate session high resolution whole-head 3D MDEFT brain scans (128 sagittal slices, 1.5 mm thickness, FOV 25.0 × 25.0 × 19.2 cm, data matrix of 256 × 256 voxels) were acquired additionally for reasons of improved localization (Lee et al., 1995; Ugurbil et al., 1993).

2.6. Data analysis

The data processing was performed using the software package LIPSIA (Lohmann et al., 2001). Functional data were corrected for slice-time acquisition differences using sinc-interpolation. In addition, the data were corrected for motion artefacts. Signal changes and baseline-drifts were removed by applying a temporal high-pass filter to remove frequencies lower than 1/48 Hz (2.5 times the length of one complete oscillation). The anatomical slices were co-registered with the full-brain scan that resided in the stereotactic coordinate system and then transformed by linear scaling to a standard size. The transformation parameters obtained from this step were subsequently applied to the functional slices so that the functional slices were also registered into the stereotactic space. The statistical analysis was based on a least squares estimation using the general linear model (GLM) for serially autocorrelated observations (Aguirre, Zarahn, & D'Esposito, 1997; Bosch, 2000; Friston, 1994; Zarahn, Aguirre, & D'Esposito, 1997). The model equation including the observation data, the design matrix, and the error term, was linearly smoothed by convolving it with a Gaussian kernel of dispersion of 4 s FWHM. The contrast between the different conditions was calculated using the *t* statistics. Subsequently, *t* values were converted to *Z* scores. As the individual functional datasets were all aligned to the same stereotactic reference space a group analysis of fMRI-data was performed by averaging individual *Z* maps. The average SPM was multiplied by a SPM correction factor of square root of the current number of subjects (*N* = 14) (Bosch, 2000). For the purpose of illustration, averaged data were superimposed onto one normalized 3D MDEFT standard volume (see Fig. 2).

To identify differential brain systems subserving passive perception and active rehearsal of intonation contour we modelled the hemodynamic response for the time window during stimulus presentation ('passive perception') and mental task performance (active reproduction) separately.

3. Results

In this section we report brain responses collected whilst volunteers were listening to speech stimuli (A, B, and C), or rehearsed sentence intonation (D). Fig. 2 and Tables 1–4 display the main results of the contrasts between conditions.

3.1. Normal speech vs. degraded speech

This contrast yielded a pattern of functional brain activation which is almost identical to the results described by Meyer et al. (2002). Listening to normal sentences involves the left IFG (pars triangularis), large parts of the left superior temporal gyrus (STG) and superior temporal sulcus (STS), the left temporo-occipital-parietal (TOP) junction area, and yields only small right hemisphere responses in the primary auditory areas (cf. Fig. 2A and Table 2). When compared to normal sentences degraded speech produced stronger functional activity in the left and right posterior superior temporal region, particularly in the planum temporale. Further areas that show stronger responses to degraded speech included the left middle frontal gyrus (MFG), the inferior parietal lobe (IPL), the right Rolandic operculum (ventral premotor cortex), and the deep frontal operculum bilaterally. Additionally the basal ganglia (Putamen, Caudate Nucleus) turned out to be more strongly involved whilst participants heard pure intonation contour.

3.2. Normal speech vs. flattened speech

Several areas showed stronger activation for flattened speech relative to normal speech including large portions of the right posterior STG and the planum temporale bilaterally (cf. Fig. 2B and Table 2). Further activation was found in the left MFG and in the left IPL.

3.3. Flattened speech vs. degraded speech

Like normal speech hearing flattened speech relative to degraded speech revealed stronger activity in the left IFG (pars triangularis) and along the entire STG/STS bilaterally, and in the left temporo-occipital-parietal (TOP) junction area, (cf. Fig. 2C, Table 3). Analogous to contrast Fig. 2A the left and right frontal operculum, the right Rolandic operculum, the right planum parietale, and the basal ganglia responded more strongly to degraded relative to flattened speech.

3.4. Rehearsal: Intonation vs. flattened speech

Fig. 2D and Table 4 shows that active rehearsal of sentence intonation involved areas only in the left

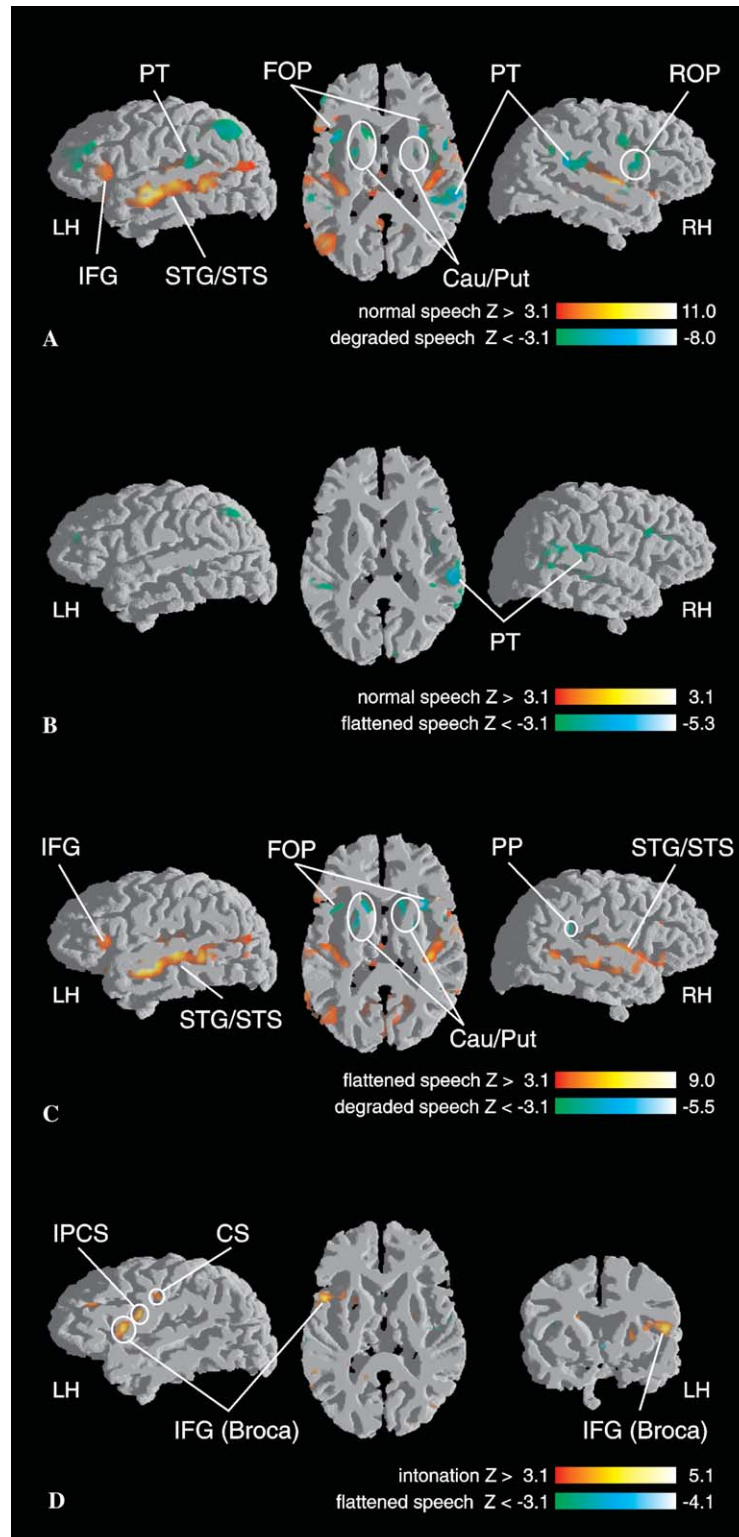


Fig. 2. Views of direct comparison between conditions whilst subjects heard normal and degraded speech (A), normal and flattened speech (B), flattened and degraded speech (C), and whilst subjects rehearsed sentence intonation compared to rehearsal of flattened speech (D). Functional inter-subject activation ($N = 14$) is plotted in neurological convention on parasagittal and horizontal slices intersecting the peri-sylvian cortex. All figures display significant brain responses ($Z \geq 3.10$, α -level 0.001) superimposed onto a normalised white-matter segmented 3D reference brain. Thus, the brain's white matter is separated from gray matter so that the cortical layers (the outermost 3–5 mm) are removed. IFG, inferior frontal gyrus; IPCS, inferior precentral sulcus; CS, central sulcus; FOP, fronto-opercular cortex; ROP, Rolandic operculum; STG, superior temporal gyrus; STS, superior temporal sulcus; PT, planum temporale; PP, planum parietale; Cau, caudate nucleus; Put, putamen.

Table 1
Normal speech vs. degraded speech

Location	BA	Left hemisphere				Right hemisphere			
		Z score	x	y	z	Z score	x	y	z
Normal speech > degraded speech									
IFG	45	7.1	-47	25	5	—	—	—	—
IFG	47	—	—	—	—	4.5	40	19	-10
mid./post. STG/STS	22/21	10.3	-53	-11	-6	9.3	51	-8	0
Tha	—	4.0	-11	-32	7	—	—	—	—
post. GC	31	4.75	-4	-51	30	—	—	—	—
Degraded speech > normal speech									
MFG	46	-5.8	-45	43	10	—	—	—	—
FOP	—	-5.7	-39	13	9	-5.9	26	16	7
ROP	44/6	—	—	—	—	-5.4	44	6	10
PrCG	6	-4.6	-53	2	15	—	—	—	—
mid. CG	—	—	—	—	—	-4.06	0	-24	29
post. STG/PT	42/22	-4.4	-51	-33	18	-7.5	56	-34	21
IPL	40/19	-4.4	-29	-48	40	-3.7	37	-40	41

In this table and in Tables 2–4 results of direct comparison of different conditions are listed. Z scores indicate the magnitude of statistical significance. Localization is based on stereotactic coordinates (Talairach & Tournoux, 1988). These coordinates refer to the location of maximal activation indicated by the Z score in a particular anatomical structure. Distances are relative, to the intercommissural (AC–PC) line in the horizontal (x), anterior–posterior (y), and vertical (z) directions. Functional activation was thresholded at $|Z| \geq 3.1$. The table only lists activation clusters exceeding a minimal size of 100 voxels. IFG, inferior frontal gyrus; MFG, mid frontal gyrus; aINS, anterior insula; FOP, frontal operculum; ROP, Rolandic operculum; PrCG, precentral gyrus; IPCS, inferior precentral sulcus; CS, central sulcus (Rolandic fissure); STG, superior temporal gyrus; STS, temporal sulcus; MTG, middle temporal gyrus; PT, planum temporale; PP, planum parietal; CG, cingulate gyrus; IPL, inferior parietal lobe; SPL, superior parietal lobe; Tha, Thalamus; Cau, caudate head; Put, putamen.

Table 2
Flattened pitch vs. normal speech

Location	BA	Left hemisphere				Right hemisphere			
		Z score	x	y	z	Z score	x	y	z
Flattened pitch > normal speech									
MFG	46	-3.9	-45	43	11	—	—	—	—
IFG	45	—	—	—	—	-4.4	39	23	8
ROP	44/6	—	—	—	—	-3.7	41	-1	6
mid STG	42	—	—	—	—	-4.8	50	-28	16
post. STG	42/22	-4.5	-58	-33	15	-4.1	38	-39	15
PP	40	—	—	—	—	-4.7	56	-42	23
IPL	40/19	-4.0	-30	-48	41	-3.9	34	-58	43
SPL	40/7	-3.8	-51	-54	48	—	—	—	—

Functional activation indicated separately for contrasts between conditions. For explanations see Table 1.

hemisphere, namely in Broca's area, in the ventral pre-motor and presumably motor cortex. Plausibly, rehearsing flattened speech did not produce any significant activity.

4. Discussion

The study examined whether we can identify brain regions particularly mediating the processing of intonational information. Therefore we delivered to participants sentences that either carried segmental cues but lacked dynamic pitch variations (*flattened speech*), solely suprasegmental cues (*degraded speech*), or both seg-

mental and suprasegmental information (normal speech).

Generally, we found auditory speech stimuli to activate brain areas including the bilateral peri-sylvian cortex. However, the different conditions produced differential activation patterns pointing to distinct brain regions which may subserve differential aspects of speech processing. Processing both normal speech and flattened speech which comprise proper lexical and syntactic information corresponded to brain activation in the left IFG, the bilateral STG/STS, and the left TOP junction area (2A and 2C). The left IFG and the STG have been comprehensively described as mediating normal sentence comprehension in the auditory domain

Table 3
Flattened pitch vs. degraded speech

Location	BA	Left hemisphere				Right hemisphere			
		Z score	x	y	z	Z score	x	y	z
Flattened pitch > degraded speech									
IFG	47	—	—	—	—	4.0	35	30	–11
IFG	45	5.4	–51	24	5	—	—	—	—
ant. STG/MTG	22/21	8.5	–53	–11	–6	9.0	51	–7	–1
CS	4/3	—	—	—	—	3.9	56	–8	23
post. MTG	21	—	—	—	—	4.8	55	–49	9
post. STS	39	4.5	–42	–67	24	—	—	—	—
post. CG	31	—	—	—	—	3.8	3	–52	24
Tha	—	4.1	–15	–30	3	—	—	—	—
Degraded speech > flattened pitch									
aINS/FOP	—	–3.9	–43	13	8	–5.4	26	16	7
ROP	44/6	—	—	—	—	–3.6	43	5	10
Cau	—	–4.9	–17	14	9	–4.5	14	13	5
Put	—	—	—	—	—	–5.2	18	5	8
post. STG/PP	42/22	—	—	—	—	–4.2	54	–35	20

Functional activation indicated separately for contrasts between conditions. For explanations see Table 2.

Table 4
Pitch variation speech vs. flattened speech

Location	BA	Left hemisphere				Right hemisphere			
		Z score	x	y	z	Z score	x	y	z
Pitch variation > flattened speech									
IFG	44	4.7	–46	9	11	—	—	—	—
IPCS	6	4.69	–52	0	20	—	—	—	—
CS	4	4.0	–46	–14	32	—	—	—	—
MFG	9	4.53	–37	33	26	4.45	28	30	29

Functional activation during rehearsal interval indicated separately for contrasts between sentence stimuli with and without dynamic pitch variation. For explanations see Table 2.

(Friederici, 2002; Friederici & Alter, this volume; Kaan & Swaab, 2002). We also found stronger responses to normal and flattened speech in left and right STG/STS (2A and C). Whilst the left STG/STS has been associated with a ‘pathway for speech comprehension’ (Binder et al., 2000; Scott, Blank, Rosen, & Wise, 2000) the stronger right hemisphere activity for normal and flattened speech demands an explanation. Degraded speech comprises attenuated spectral complexity. Therefore we argue that the right STG/STS activity accounts for processing spectral information. Additionally both normal and flattened speech activated the left TOP junction area. This finding accords with a recent model by Hickok & Pöppel (2000) who proposed that the TOP junction area constitutes an auditory–conceptual interface. In contrast to degraded speech, processing normal and flattened speech requires access to the mental lexicon. In terms of the present study, activity in the TOP junction area reflects access to meaning-based representations during sentence comprehension. Additionally to normal sentence comprehension which more strongly recruited left brain regions the study revealed more salient brain responses in right peri-sylvian areas whilst the

volunteers heard phonetically incongruent sentences either lacking dynamic pitch contour (flattened speech) or comprising melodic intonation (degraded speech). In a cross-linguistic study which examined the temporal integration of speech prosody in native speakers of Chinese and English Gandour and colleagues demonstrated that local pitch contours associated with tones are processed in the left hemisphere whereas global pitch contours representing intonation are processed in the right hemisphere (Gandour et al., 2003). The functional rightward asymmetry was even more prominent when native speakers of English heard pure intonation contours (hummed versions of Chinese sentences). Furthermore the authors propose that native speakers of intonational languages (like English) are more able to discern pitch contours which span over longer temporal domains whereas native speakers of tonal languages (like Chinese and Thai) are more primed to picking up pitch contours which span over shorter temporal domains. These observations in compliance with our present data partially agree with several recent neuro-cognitive models that sketched a bilateral cerebral implementation of speech perception (Friederici, 2002;

Hickok & Pöppel, 2000) with the right hemisphere playing a prominent role in processing intonation (Dogil et al., 2002; Friederici & Alter, this volume; Pöppel, 2003). Building on the results of the present study we can argue against recent views proposing a universal right hemisphere superiority for propositional prosody per se (Weintraub & Mesulam, 1981), but we can support the view that bilaterally distributed mechanisms may constitute the ability to decode the melodic line of speech (Pell & Baum, 1997). However, it appears that hemispheric specialization is dependent on the 'prosodic needs' of a particular language (Gandour et al., 2002, 2003) and to what extent spectral and temporal parameters shape the linguistic and emotional interpretation of the received speech (Lakshminarayanan et al., 2003).⁴

Nonetheless, the present finding needs a comprehensive discussion: Normal speech signals, (i.e., a sentence) comprise rapidly changing (formant transitions, high frequencies) and slowly changing (syllable-seized changing intonation, fundamental frequency) acoustic features. In flattened speech we primarily removed slow modulations of the pitch contour resulting in a monotonous sounding utterance. Degraded speech primarily lacks rapidly changing cues. Brain responses to flattened and degraded speech were not identical but overlapped in the posterior STG bilaterally, particularly in the planum temporale (PT) (2A and B). Functionally, the PT corresponds to auditory association cortices that are important in higher order processing of auditory language input (Foundas, Leonard, Gilmore, Fennel, & Heilman, 1994). Morphologically, right-handers demonstrate a significant volumetric PT leftward asymmetry (Foundas, Leonard, & Hanna-Pladdy, 2002; Steinmetz, Volkman, Jäncke, & Freund, 1991). However, it appears that this observed leftward asymmetry reflects a difference in morphology rather than size of the left and right PT (Westbury, Zatorre, & Evans, 1999). There has been an initial tendency to consider the differential morphologic characteristics of the left PT as the likely neuroanatomical substrate of language. Meanwhile recent findings give argument in favour of the left PT mediating the analysis of rapidly changing cues relative to language or even phonetic analysis per se (Celsis et al., 1999; Jäncke et al., 2001, 2002). Complementary, a functional rightward asymmetry to extended relative to rapid formant transitions has been proposed (Belin et al., 1998). Corroborating evidence stems from several

studies that implicate a functional designation of computing harmonic information to the right hemisphere and temporal cues to the left hemisphere (Johnsrude, Zatorre, Milner, & Evans, 1997; Zatorre & Belin, 2001; Zatorre, Belin, & Penhune, 2002). It appears that this neurofunctional assignment holds more for more inferolaterally situated parts of the STG/STS rather than the PT. Since both flattened and degraded speech activated both the left and right PT our data do not fully support the notion of the left PT processing only rapidly changing cues and the right PT solely processing slowly changing cues. The present data rather support a view which suggests the left and right PT as governing the integration of rapidly and slowly changing acoustic cues during speech comprehension. Thus, the starker responses to flattened or degraded speech are caused by the phonetically degraded inflowing segmental and suprasegmental units and reflect detection and compensation of incongruent acoustic information. Our view of the PT constituting an integrating mechanism is consistent with a recent review article describing the PT as a 'computational hub' which matches spectrotemporal acoustic patterns (Griffiths & Warren, 2002). The right planum parietale (PP) also brought on an increase in activity for impoverished speech, i.e., when degraded and flattened speech is compared to normal speech. However, based on the current data it cannot be concluded as to whether attentional demands account for increase in activity observed in the PP (Jäncke et al., 2001, 1994) or as to whether the right posterior Sylvian Fissure is generally more concerned with slow prosodic modulations as available in degraded speech.

In addition to activity in the PT and PP, we found an ensemble of bilateral fronto-opercular areas including basal ganglia and the right ROP more strongly responding to degraded speech when compared to normal and flattened speech (2A and 2C) which is a replication of the results described by Meyer et al. (2002). Briefly summarized, we suggest that activity in the deep fronto-opercular cortex reflects an interaction between processing unintelligible speech and related effort. We will be discussing the functions of the different brain structures in turn, starting with the basal ganglia. Furthermore, we will be discussing the role the intrasylvian cortex may play during the perception of degraded speech and its implication for speech motor control.

The data exposed an involvement of basal ganglia (Putamen, Caudate Nucleus) in processing degraded relative to normal and flattened speech (see Figs. 2A and C). In combination with the ventral striatum the Putamen and the Nucleus Caudatus form the neostriatum (NS) (Murdoch, 2001). The NS is considered to be a 'multi laned pathway' which integrates diverse input from the entire cerebral cortex and projects back to cortical areas via thalamic regions ('basal ganglia-thalamicocortical circuits'). Traditionally, there is an

⁴ Several recent brain imaging studies tried to identify the brain regions which are associated with processing emotionally toned speech. Whilst some studies report stronger right hemisphere responses to emotional tone available in spoken utterances (Buchanan et al., 2000; George et al., 1996; Stiller et al., 1997; Wildgruber, Pihan, Erb, Ackermann, & Grodd, 2002), there are other studies which argue against a functional rightward asymmetry for emotional prosody (Jäncke, Buchanan, Lutz, & Shah, 2001; Kotz et al., 2003).

understanding of the basal ganglia playing a role in sequencing and generating motor functions, particularly in articulation since the Caudate Nucleus projects to the inferior ventral premotor cortex which governs supralaryngeal functions of tongue, lips, and jaws (Pickett, Kuniholm, Protopapas, Friedman, & Lieberman, 1998). Wise, Murray, & Gerfen (1996) outlined a mechanism for the basal ganglia's proposed role in rule potentiation. According to this mechanism the basal ganglia activation increase when to alter routine behaviour is needed. However, the basal ganglia have also been associated more specifically with functions related to generating, encoding, modifying, and transferring linguistic information (Crosson, 1992). Both views are principally consistent with the present data. With respect to the first, the presentation of pure sentence melody might mean an alteration of normal processing routines which is reflected by increased basal ganglia activity. The second view receives some support from clinical studies which examined sentence comprehension in patients suffering from brain lesions and from Parkinson's disease, respectively (Friederici, Kotz, Werheid, Hein, & von Cramon, 2003a; Friederici, von Cramon, & Kotz, 1999). This data suggests that deficient function of the basal ganglia affects integrational processes which occur at a late stage during sentence comprehension. Here, we observed that the basal ganglia brought on stronger activity for pure prosodic contour relative to normal and flattened speech comprising syntactic information which might reflect difficulties in integrating segmental and suprasegmental units. Recent lesion studies also speak to the issue of basal ganglia involvement in prosody. These studies provide consistent evidence that damage to the basal ganglia causes prosodic comprehension deficits (Kempler & Van Lancker, 2002; Lloyd, 1999), particularly when patients were presented with nonsense sentences (Pell, 1996) or degraded speech (Speedie, Brake, Folstein, Bowers, & Heilman, 1990). Even though these observations are not completely in harmony they might be reconciled by a more general view which proposes that the basal ganglia perform temporal binding necessary for any kind of integrational sequencing, that is timing and ordinality in higher cognition including language functions (Pickett et al., 1998). The authors purport that sequencing can be considered as an underlying device essential for syntax, prosody, and cognition. To support this view the authors report a case study of a patient with damage to Putamen and Caudate Nucleus who demonstrates deficient syntactic processing and disturbed prosodic production. An alternative explanation for the stronger basal ganglia activation in degraded speech also considers the stronger activity in the right ROP whilst participants passively perceived intonation contour (2A). Meyer et al. (2002) recently reported selective responses to intonation contour in this particular area combined with stronger basal

ganglia activity. Several recent neuroimaging studies associate the ROP with covert speech production and both covert and overt singing (Perry et al., 1999; Riecker, Ackermann, Wildgruber, Dogil, & Grodd, 2000; Wildgruber, Ackermann, Klose, Kardatzki, & Grodd, 1996). Here, subjects were asked to carefully listen to sentence melodies and to silently rehearse these melodies for several seconds. Plausibly, this representation induced sub-articulatory processes which might account for the right premotor and basal ganglia activation. Corroborating view comes from a recent brain imaging study which also reported premotor activation whilst participants processed degraded speech which involved covert speech processing and consequently strategies inducing sub-articulation (Kiehl, Laurens, & Liddle, 2002).

From the point of an integration of language and motor networks it is notable that a network of areas which have been identified to process *auditory input*, has also been described to support *producing auditory output* (Ackermann & Riecker, this volume; Dogil et al., 2002; Indefrey et al., 2001; Indefrey, Hellwig, Herzog, Seitz, & Hagoort, this volume). Clinical and neuroimaging studies suggest that a similar network subserves the production and perception of prosody (Dogil et al., 2002; Pell & Baum, 1997). However, there seems no single area which can be considered as an accommodation of the prosody generator. It rather seems that differential brain areas contribute to the production of prosodically modulated speech dependent on the prosodic frame length (Behrens, 1989; Dogil et al., 2002) with the left hemisphere subserving prosodic processing on the syllable level and the right hemisphere supporting processing on the level of intonational phrases.

In a recent fMRI study Ackermann and Riecker demonstrated that (overt) speech tasks activated the left anterior insula whilst (overt) reproduction of nonlyrical tune evoked stronger responses in the right anterior insular cortex (Ackermann & Riecker, this volume) which led the authors to suggest a distributional operational mode whereby the production of segmental information recruits the left anterior insula and the production of intonation contours engages the contralateral homologue. Even though Ackermann and Riecker conclude that activation of the intrasylvian cortex is tied to overt task performance, probably reflecting the coordination of the muscles engaged in articulation and phonation, there is also evidence in favour of the (right) anterior insula as mediating processing rather than production of auditory and speech signals (Ackermann et al., 2001; Benson et al., 2001; Meyer et al., 2002; Plante, Creusere, & Sabin, 2002; Wong et al., 2002). Interestingly, these studies report exceptionally strong responses when subjects heard nonspeech sounds which had been derived from speech signals and preserved phonetic and

prosodic information also available in normal speech sounds. To reconcile these observations with the view of the anterior insula as mediator of pure articulatory functions as put forward by Dronkers (1996) we favour the suggestion to consider the anterior insula as the location of common pre-articulatory/articulatory neurons subserving both speech perception and production (Benson et al., 2001), and possibly supporting functions of language and motor integration.

Furthermore, we observed a shift from right to left frontal areas for the post stimulus presentation interval (see Fig. 2D). Due to the instructions participants were to actively rehearse the sentence intonation which appears to correspond to a selective activity in Broca's area and in the caudally adjacent (pre-)motor cortex (ROP and central sulcus). A similar shift of functional lateralization was recently observed by Plante et al. (2002) who discovered larger intensity for rehearsal of lingual melody in right premotor areas whereas the contralateral cortex was more strongly activated in rehearsing normal sentences. Supporting evidence springs from observations which tied the left inferior frontal cortex, particularly the lateral convexity of Broca's area, to verbal rehearsal (Paulesu et al., 1993). Thus, it seems plausible to assume that verbal information available in normal speech automatically activated the left IFG in the active rehearsal condition since this region is considered to mediate sensory-motor and audio-motor integration systems which are crucial in rehearsing both segmental and suprasegmental cues (Dronkers, 1996; Hickok & Pöppel, 2000). Neuroanatomical evidence also lends credence to this view as the bilateral inferior ventral premotor cortex represents supralaryngeal articulators (lips, tongue) which indisputably play an essential role in (sub-)articulation (Kolb & Whishaw, 1995).

Based on these converging observations we finally propose a neural network consisting of brain areas which are known to reside language (fronto-opercular region), premotor functions (ROP), and sequencing functions (basal ganglia and inferior ventral premotor cortex) to coincidentally support the processing of intonation contour in spoken sentences. However, further neuroimaging research in compliance with computational modeling is required to reveal the neural underpinnings of language and motor integration in more detail.

5. Concluding remarks

In this paper we have demonstrated that [1] areas in the right rather than left peri-sylvian region are more sensitive to phonetically incongruent speech signals, that [2] the pSTG bilaterally, particularly the PT constitutes the integration of slowly and rapidly changing acoustic cues during comprehension of spoken language, that [3]

processing speech melody per se recruits a bilateral circuit involving the bilateral fronto-opercular cortex, basal ganglia, and the right premotor cortex. Finally, this fMRI study showed [4] a right fronto-lateral network for processing and a left fronto-lateral network for producing prosodic information. In sum, the data clearly adds weight to recent neurocognitive models which propose a differential contribution of right and left brain areas to prosodic processing rather than a sole right hemispheric responsibility.

Acknowledgments

The authors wish to thank Adam McNamara and two anonymous reviewers for helpful comments on the manuscript. The work was supported by the Leibniz Science Prize awarded to Angela Friederici, a grant from the German Research Foundation (FR 519/17-3) awarded to Angela Friederici and Kai Alter, and a grant from the Human Frontier Science Program (HFSP RGP 5300/2002-C102) awarded to Kai Alter.

References

- Ackermann, H. & Riecker, A. (this volume). The contribution of the insula to motor aspects of speech production: A review and a hypothesis. *Brain and Language*.
- Ackermann, H., Riecker, A., Mathiak, K., Erb, M., 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*, *12*, 4087–4092.
- Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1997). Empirical analysis of BOLD-fMRI statistics. II. Spatially smoothed data collected under null-hypothesis and experimental conditions. *NeuroImage*, *5*, 199–212.
- Behrens, S. (1989). Characterizing sentence intonation in a right-hemisphere damaged population. *Brain and Language*, *37*, 181–200.
- Belin, P., Zilbovicius, M., Crozier, S., Thivard, L., Fontaine, A., Masure, M., & Samson, Y. (1998). Lateralization of speech and auditory temporal processing. *Journal of Cognitive Neuroscience*, *10*, 536–540.
- Benson, R. R., Whalen, D. H., Richardson, M., Swainson, B., Clark, V. P., Lai, S., & Liberman, A. M. (2001). Parametrically dissociating speech and nonspeech perception in the brain using fMRI. *Brain and Language*, *78*, 364–396.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Springer, J. A., Kaufman, J. N., & Possing, E. T. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex*, *10*, 512–528.
- Bookheimer, S. (2002). Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Reviews of Neuroscience*, *25*, 151–188.
- Bosch, V. (2000). Statistical analysis of multi-subject fMRI data: The assessment of focal activations. *Journal of Magnetic Resonance Imaging*, *11*, 61–64.
- Buchanan, T., Lutz, K., Mirzazade, S., Specht, K., Shah, N. J., Zilles, K., & Jäncke, L. (2000). Recognition of emotional prosody and verbal components of spoken language: An fMRI study. *Cognitive Brain Research*, *9*, 227–238.

- Celsis, P., Boulanouar, K., Doyon, B., Ranjeva, J. P., Berry, I., Nespoulous, J. L., & Chollet, F. (1999). Differential fMRI responses in the left posterior superior temporal gyrus and left supramarginal gyrus to habituation and change detection in syllables and tones. *NeuroImage*, 9, 135–144.
- Crosson, B. (1992). *Subcortical functions in language and memory*. New York: Guilford Press.
- D'Esposito, M., Zarahn, E., & Aguirre, G. K. (1999). Event-related functional MRI: Implications for cognitive psychology. *Psychological Bulletin*, 125, 155–164.
- Dogil, G., Ackermann, H., Grodd, W., Haider, H., Kamp, H., Mayer, J., Riecker, A., & Wildgruber, D. (2002). The speaking brain: A tutorial introduction to fMRI experiments in the production of speech, prosody, and syntax. *The Journal of Neurolinguistics*, 15, 59–90.
- Dronkers, N. (1996). A new brain region for coordinating speech articulation. *Nature*, 384, 159–161.
- Foundas, A. L., Leonard, C. M., Gilmore, R., Fennel, E., & Heilman, K. M. (1994). Planum temporale asymmetry and language dominance. *Neuropsychologia*, 32, 1225–1231.
- Foundas, A. L., Leonard, C. M., & Hanna-Pladdy, B. (2002). Variability in the anatomy of the planum temporale and posterior ascending ramus: Do right- and left handers differ? *Brain and Language*, 83, 403–424.
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Science*, 6, 78–84.
- Friederici, A. D. & Alter, K. (this volume). Lateralization of auditory language functions: A dynamic dual pathway model. *Brain and Language*.
- Friederici, A. D., Kotz, S. A., Werheid, K., Hein, G., & von Cramon, D. Y. (2003a). Syntactic comprehension in Parkinson's disease: Investigating early automatic and late integrational processes using ERPs. *Neuropsychology*, 17, 133–142.
- Friederici, A. D., Rüschemeyer, S., Hahne, A., & Fiebach, C. (2003b). The role of left inferior frontal and superior temporal cortex in sentence comprehension: Localizing syntactic and semantic processes. *Cerebral Cortex*, 13, 170–177.
- Friederici, A. D., von Cramon, D. Y., & Kotz, S. A. (1999). Language related brain potentials in patients with cortical and subcortical left hemisphere lesions. *Brain*, 122, 1033–1047.
- Friston, K. J. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2, 189–210.
- Gandour, J., Dziedzic, M., Wong, D., Lowe, M., Tong, Y., Hsieh, L., Sathamnuwong, N., & Lurito, J. (2003). Temporal integration of speech prosody is shaped by language experience: An fMRI study. *Brain and Language*, 84, 318–336.
- Gandour, J., Wong, D., Lowe, M., Dziedzic, M., Sathamnuwong, N., Tong, Y., & Li, X. (2002). A cross-linguistic fMRI study of spectral and temporal cues underlying phonological processing. *Journal of Cognitive Neuroscience*, 14, 1076–1087.
- George, M. S., Priti, P. I., Rosinky, N., Ketter, T. A., Kimbrell, T. A., Heilman, K. M., Herscovitch, P., & Post, R. (1996). Understanding emotional prosody activates right hemisphere regions. *Archives of Neurology*, 53, 665–670.
- Griffiths, T. D., & Warren, J. D. (2002). The planum temporale as a computational hub. *Trends in Neuroscience*, 25, 348–353.
- Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Science*, 4, 131–138.
- Indefrey, P., Brown, C. M., Hellwig, F., Amunts, K., Herzog, H., Seitz, R. J., & Hagoort, P. (2001). A neural correlate of syntactic encoding during speech production. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 5933–5936.
- Indefrey, P., Hellwig, F., Herzog, H., Seitz, R. J. & Hagoort, P. (this volume). Neural responses to the production and comprehension of syntax in identical utterances. *Brain and Language*.
- Jäncke, L., Buchanan, T., Lutz, K., & Shah, N. J. S. (2001). Focused and non-focused attention in verbal and emotional dichotic listening: An fMRI study. *Brain and Language*, 78, 349–363.
- Jäncke, L., Schlaug, G., Huang, X., & Steinmetz, H. (1994). Asymmetry of the planum parietale. *NeuroReport*, 5, 1161–1163.
- Jäncke, L., Wüstenberg, T., Scheich, H., & Heinze, H. (2002). Phonetic perception and the temporal cortex. *NeuroImage*, 15, 733–746.
- Johnsrude, I. S., Penhune, B., & Zatorre, R. J. (2000). Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain*, 123, 155–163.
- Johnsrude, I. S., Zatorre, R. J., Milner, B. A., & Evans, A. C. (1997). Left-hemisphere specialization for the processing of acoustic transients. *NeuroReport*, 8, 1761–1765.
- Kaan, E., & Swaab, T. Y. (2002). The brain circuitry of syntactic comprehension. *Trends in Cognitive Science*, 6, 350–356.
- Kempler, D., & Van Lancker, D. (2002). Effect of speech task on intelligibility in Dysarthria: A case study of Parkinson's disease. *Brain and Language*, 80, 449–464.
- Kiehl, K. A., Laurens, K. R., & Liddle, P. F. (2002). Reading anomalous sentences: An event-related fMRI-study of semantic processing. *NeuroImage*, 17, 842–850.
- Kolb, B., & Wishaw, I. Q. (1995). *Fundamentals of human neuropsychology*. New York: Freeman.
- Kotz, S. A., Meyer, M., Alter, K., Besson, M., von Cramon, D. Y., & Friederici, A. D. (2003). On the lateralization of emotional prosody: An event-related functional MR investigation. *Brain and Language*, 86, 366–376.
- Lakshminarayanan, K., Shalom, D., van Wassenhove, V., Orbelo, D., Houde, J., & Poeppel, D. (2003). The effect of spectral manipulation on the identification of affective and linguistic prosody. *Brain and Language*, 84, 250–263.
- Lee, J.-H., Garwood, M., Menon, R., Adrian, G., Andersen, P., Truwit, C. L., & Ugurbil, K. (1995). High contrast and fast three-dimensional magnetic resonance imaging at high fields. *Magnetic Resonance in Medicine*, 34, 308.
- Lloyd, A. J. (1999). Comprehension of prosody in Parkinson's disease. *Cortex*, 35, 389–402.
- Lohmann, G., Müller, K., Bosch, V., Mentzel, H., Hessler, S., Chen, L., Zysset, S., & von Cramon, D. Y. (2001). Lipsia—a new software system for the evaluation of functional magnetic resonance images of the human brain. *Computerized Medical Imaging and Graphics*, 25, 449–457.
- Meyer, M., Alter, K., Friederici, A. D., Lohmann, G., & von Cramon, D. Y. (2002). Functional MRI reveals brain regions mediating slow prosodic manipulations of spoken sentences. *Human Brain Mapping*, 17, 73–88.
- Murdoch, B. E. (2001). Subcortical brain mechanism in speech and language. *Folia Phoniatrica et Logopaedica*, 53, 233–251.
- Oldfield, R. C. (1971). The assessment and analysis of handedness. *Neuropsychologia*, 9, 97–113.
- Paulesu, E., Frith, C. D., & Frackowiak, R. S. J. (1993). The neural correlates of the verbal components of working memory. *Nature*, 362, 342–345.
- Pell, M. D. (1996). On the receptive prosodic loss in Parkinson's disease. *Cortex*, 32, 693–704.
- Pell, M. D., & Baum, S. R. (1997). The ability to perceive and comprehend intonation in linguistic and affective contexts by brain-damaged adults. *Brain and Language*, 57, 80–99.
- Perry, D. W., Zatorre, R. J., Petrides, M., Alivisatos, B., Meyer, E., & Evans, A. C. (1999). Localization of cerebral activity during simple singing. *NeuroReport*, 11, 3979–3984.
- Pickett, E. R., Kuniholm, E., Protopapas, A., Friedman, J., & Lieberman, P. (1998). Selective speech motor, syntax and cognitive deficits with bilateral damage to the putamen and the head of caudate nucleus: A case study. *Neuropsychologia*, 36, 173–188.

- Plante, E., Creusere, M., & Sabin, C. (2002). Dissociating sentential prosody from sentence processing: Activation interacts with task demands. *NeuroImage*, 17, 401–410.
- Pöppel, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as 'asymmetric sampling in time'. *Speech and Communication*, 41, 245–255.
- Riecker, A., Ackermann, H., Wildgruber, D., Dogil, G., & Grodd, W. (2000). Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *NeuroReport*, 11, 1997–2000.
- Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. S. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, 123, 2400–2406.
- Sonntag, G. P., & Portele, T. (1998). PURR—a method for prosody evaluation and investigation. *Journal of Computer Speech and Language*, 12, 437–451.
- Speedie, L. J., Brake, N., Folstein, S. E., Bowers, D., & Heilman, K. M. (1990). Comprehension of prosody in Huntington's disease. *The Journal of Neurology Neurosurgery and Psychiatry*, 53, 607–610.
- Steinhauer, K. (2003). Electrophysiological correlates of prosody and punctuation. *Brain and Language*, 86, 142–164.
- Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience*, 2, 191–196.
- Steinmetz, H., Volkman, J., Jäncke, L., & Freund, H. (1991). Anatomical left-right asymmetry of language-related temporal cortex is different in left and right handers. *Annals of Neurology*, 29, 315–319.
- Stiller, D., Gaschler-Markefski, B., Baumgart, F., Schindler, F., Tempelmann, C., Heinze, H.-J., & Scheich, H. (1997). Lateralized processing of speech prosodies in the temporal cortex: A 3-T functional magnetic resonance imaging study. *Magma*, 5, 275–284.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. New York: Thieme.
- Ugurbil, K., Garwood, M., Ellermann, J., Hendrich, K., Hinke, R., Hu, X., Kim, S.-G., Menon, R., Merkle, H., Ogawa, S., & Salmi, R. (1993). Magnetic fields: Initial experiences at 4T. *Magnetic Resonance Quarterly*, 9, 259.
- Weintraub, S., & Mesulam, M.-M. (1981). Disturbances of prosody. A right-hemisphere contribution to language. *Archives of Neurology*, 38, 742–744.
- Westbury, C. F., Zatorre, R. J., & Evans, A. C. (1999). Quantifying variability in the planum temporale: A probability map. *Cerebral Cortex*, 9, 392–405.
- 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., Pihan, D., Erb, M., Ackermann, H., & Grodd, W. (2002). Dynamic brain activation during processing of emotional intonation: Influence of acoustic parameters, emotional valence, and sex. *NeuroImage*, 15, 856–869.
- Wise, S. P., Murray, E. A., & Gerfen, C. R. (1996). The frontal cortex - basal ganglia system in primates. *Critical Reviews in Neurobiology*, 10, 317–356.
- Wong, D., Pisoni, D. B., Learn, J., Gandour, J. T., Miyamoto, R. T., & Hutchins, G. D. (2002). PET imaging of differential cortical activation by monaural speech and nonspeech stimuli. *Hearing Research*, 166, 9–23.
- Zarahn, E., Aguirre, G., & D'Esposito, M. (1997). Empirical analysis of BOLD-fMRI statistics. I. Spatially smoothed data collected under null-hypothesis and experimental conditions. *NeuroImage*, 5, 179–197.
- Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, 11, 946–953.
- Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Science*, 6, 37–46.