Research Report

Segmental processing in the human auditory dorsal stream

Tino Zaehle[a,*], Eveline Geiser[a], Kai Alter[b], Lutz Jancke[a], Martin Meyer[a,c]

[a]Department of Neuropsychology, University of Zurich, Binzmühlestrasse 14/25, 8050 Zurich, Switzerland
[b]Institute of Neuroscience, Newcastle Auditory Group, Newcastle University, Newcastle NE1 4HH, UK
[c]Institute of Neuroradiology, University Hospital of Zurich, Zurich, Switzerland

ARTICLE INFO

ABSTRACT

In the present study we investigated the functional organization of sublexical auditory perception with specific respect to auditory spectro-temporal processing in speech and non-speech sounds. Participants discriminated verbal and nonverbal auditory stimuli according to either spectral or temporal acoustic features in the context of a sparse event-related functional magnetic resonance imaging (fMRI) study. Based on recent models of speech processing, we hypothesized that auditory segmental processing, as is required in the discrimination of speech and non-speech sound according to its temporal features, will lead to a specific involvement of a left-hemispheric dorsal processing network comprising the posterior portion of the inferior frontal cortex and the inferior parietal lobe. In agreement with our hypothesis results revealed significant responses in the posterior part of the inferior frontal gyrus and the parietal operculum of the left hemisphere when participants had to discriminate speech and non-speech stimuli based on subtle temporal acoustic features. In contrast, when participants had to discriminate speech and non-speech stimuli on the basis of changes in the frequency content, we observed bilateral activations along the middle temporal gyrus and superior temporal sulcus. The results of the present study demonstrate an involvement of the dorsal pathway in the segmental sublexical analysis of speech sounds as well as in the segmental acoustic analysis of non-speech sounds with analogous spectro-temporal characteristics.

1. Introduction

The cortical areas devoted to auditory speech processing are located in the middle and superior temporal gyri, superior temporal sulcus, and inferior frontal cortical areas of the left and right hemisphere. Since the seminal studies of Wernicke (1874) in the 19th century, the complementary roles of the left and right temporal lobes in auditory processing have traditionally been examined in brain-lesioned patients with behavioral assessment tools and, in the past few years, with functional and structural brain imaging methods. The majority of these studies have supported the notion of the predominant role of the left hemisphere in speech processing (Crinion et al., 2003) and point to a stronger involvement of the right hemisphere in music processing (Tervaniemi et al., 2000). In more recent times attempts have been made to identify the properties that might account for the hemispheric differences in auditory processing. Speech-related left hemisphere lateralization has been attributed to the linguistic relevance and intelligibility of the input (Scott et al., 2000). Alternatively, it has been argued that the hemispheric specialization exists as a result of asymmetries in basic auditory processing with the

left hemisphere preferentially driven by rapidly changing acoustic cues, namely formant transitions (Efron, 1963; Tallal and Piercy, 1973). More specifically, Zatorre et al. (2002) have hypothesized that asymmetries in auditory processing may be considered the developmental outcome of optimizing the processing of acoustic cues, with left auditory cortical areas being highly adept at temporal resolution and right auditory cortical areas being more amenable to spectral resolution. Along a similar line of argument, Poeppel (2001, 2003) has developed a framework in which asymmetries with respect to speech perception may be accounted for by hemispheric differences in sampling time: the left auditory areas preferentially extract information from short temporal integration windows ($\sim$40 Hz) and the right auditory areas from long integration windows ($\sim$4 Hz).

These models of functional hemisphere asymmetry in elemental auditory processing have supplied the groundwork for large-scale neurofunctional models of speech processing. Underlying these models is the assumption that there are two distinct processing streams in the auditory domain, analogous to the dual stream visual processing theory. The notion of a functional segregation in the auditory system is based on seminal findings from experimental work in the rhesus monkey, demonstrating two parallel processing streams specifically devoted to object recognition and object localization (Rauschecker, 1998b; Rauschecker and Tian, 2000). With respect to human speech perception, Scott and Wise (2004) have outlined a functional neuroanatomical framework along these lines: an anterior processing stream running anterolaterally to the primary auditory cortex is involved in sound-to-meaning mapping and a posterior stream located in the temporoparietal junction subserves the mapping of speech sounds to motor representations of articulations (Scott and Wise, 2004). The model on functional anatomy of speech elaborated by Hickok and Poeppel (2004, 2007) is also based on a distinction between two processing streams, constituting a functional dorsal–ventral partitioning. The ventral stream, projecting ventro-laterally and involving the superior temporal sulcus (STS) bilaterally, the posterior middle temporal gyrus (MTG) and the inferior temporal gyrus (ITG), mediates the mapping of sound-to-meaning. The dorsal stream, projecting dorso-posteriorly towards the inferior parietal lobe and frontal regions of the left hemisphere, supports the sound-to-articulatory-based representations. This dorsal pathway is taken to play an important role in sublexical speech perception (Hickok and Poeppel, 2000). The authors hypothesized that the performance of sublexical tasks specifically engages the dorsal processing path connecting left inferior frontal and left inferior parietal cortex as part of a network for audio-motor integration. Their reasoning is based on lesion data (Caplan et al., 1995) as well as on neuroimaging studies (Burton et al., 2000) that showed dissociations between auditory comprehension tasks and overt sublexical phoneme tasks, the former task being associated with the bilateral temporal cortex and the latter task with the inferior frontal cortex. There is strong evidence that the left inferior frontal gyrus (IFG) is specifically engaged in sublexical tasks that require explicit segmentation of the auditory signal (Burton, 2001; Burton and Small, 2006). Furthermore, activation of the left IFG has been associated with the discrimination of dynamic auditory cues in speech

(Binder et al., 1997; Poldrack et al., 2001; Zatorre et al., 1992) and non-speech sounds (Fiez et al., 1996; Joanisse and Gati, 2003), indicating a general role of the left IFG in the processing of transient acoustic information.

In the present study we investigated the neurofunctional representation of sublexical auditory processing by comparing hemodynamic responses during the discrimination of speech and non-speech sounds based on temporal or spectral variations. Participants performed a discrimination task on four experimental conditions: the discrimination of non-speech sounds based on temporal (NS_{temp}) and on spectral features (NS_{spect}), and the discrimination of speech sounds based on temporal (S_{temp}) and on spectral features (S_{spect}) in the context of a sparse fMRI-design. Statistical analysis evaluated the significance of the main effects linguistic stimulus category (LSC) and acoustic discrimination criterion (ADC). The effect of LSC was evaluated by comparing the processing of the speech sounds with the processing of the non-speech sounds independent of the temporal or spectral variations. The effect of ADC was analyzed by comparing the processing of both speech and non-speech sounds based on temporal variations with the processing of these sounds based on spectral variations.

Based on the models of speech processing outlined above, we hypothesized that segmental auditory processing occurring when sublexical speech discrimination is based on temporal acoustic features, will specifically involve the left-hemispheric dorsal processing network. Furthermore, we hypothesized that the dorsal processing stream is not driven specifically by speech but is also dedicated to the discrimination of non-speech sounds based on temporal features.

2. Results

2.1. Behavioral results

Behavioral data were successfully collected from all participants. Mean accuracy data (percent correct responses) for NS_{temp} (85.3%, SD 6.7), NS_{spect} (87.2%, SD 7.5), S_{temp} (86.9%, SD 7.5) and S_{spect} (87.8%, SD 6.6) were subjected to a 2×2 ANOVA with the factors LSC (speech/non-speech) and ADC (temporal/spectral). The analysis revealed no significant main effects of either the factor LSC ($F(1,15)=0.1$, $p=0.9$) nor of the factor ADC ($F(1,15)=0.2$, $p=0.7$) demonstrating that the mean accuracies did not differ between either the speech and non-speech condition, or between the spectral and temporal discrimination.

2.2. Functional MRI results

In a first step, each of the four stimulus conditions was compared with the resting baseline condition. As illustrated in Fig. 1, the analysis revealed significant activation along the bilateral superior temporal plane for all four conditions. Subsequent ROI analysis on anatomically predefined regions (bilateral planum polare (PP), Heschl's gyrus (HG), and planum temporale (PT)) showed a general functional leftward asymmetry for all conditions and regions (cf. Fig. 2). Separate $2 \times 2 \times 2$ ANOVAs with the factors LSC (speech/non-speech), ADC (temporal/spectral), and hemisphere (left/right) revealed a significant main effect of the factor hemisphere for PP ($F(1,15)=11.12$,
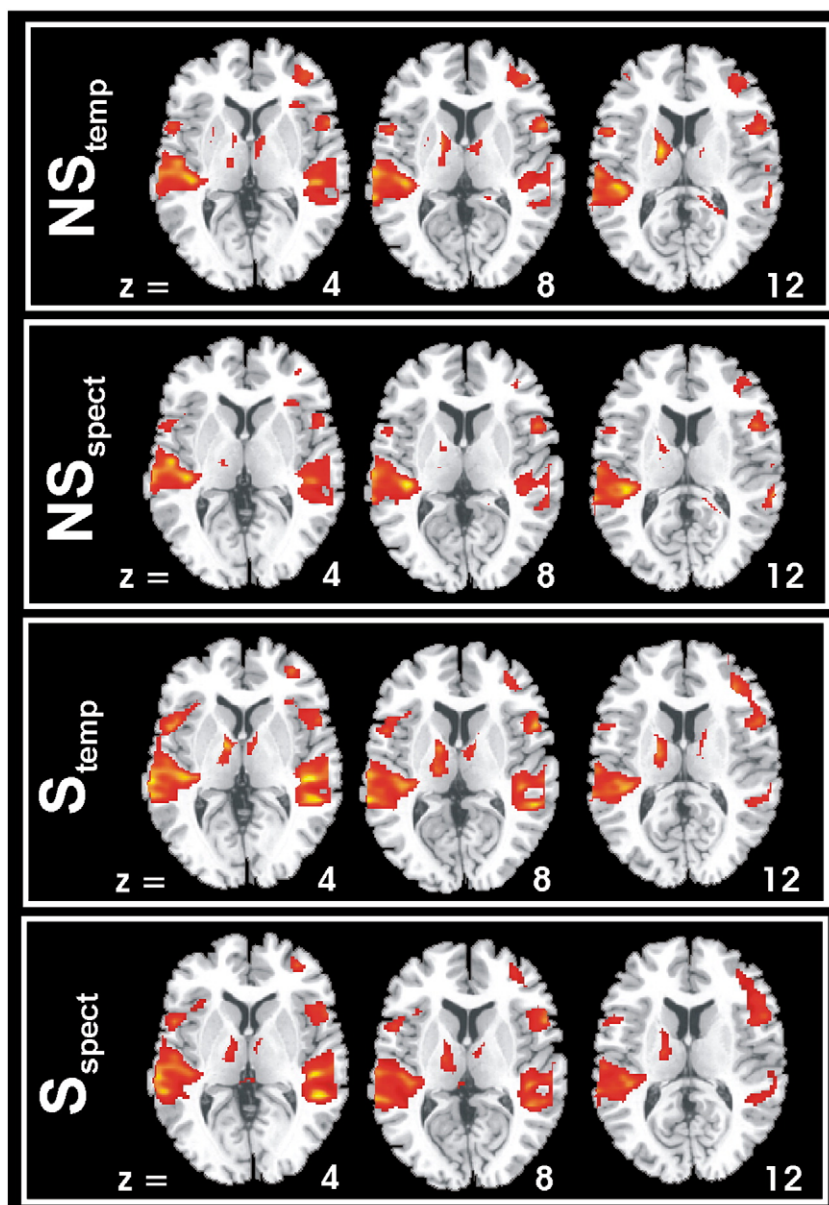
Fig. 1 – Activation pattern for the comparison for each condition with the resting baseline. Results are superimposed onto horizontal sections of a standard anatomical template. Data are plotted in neurological convention.
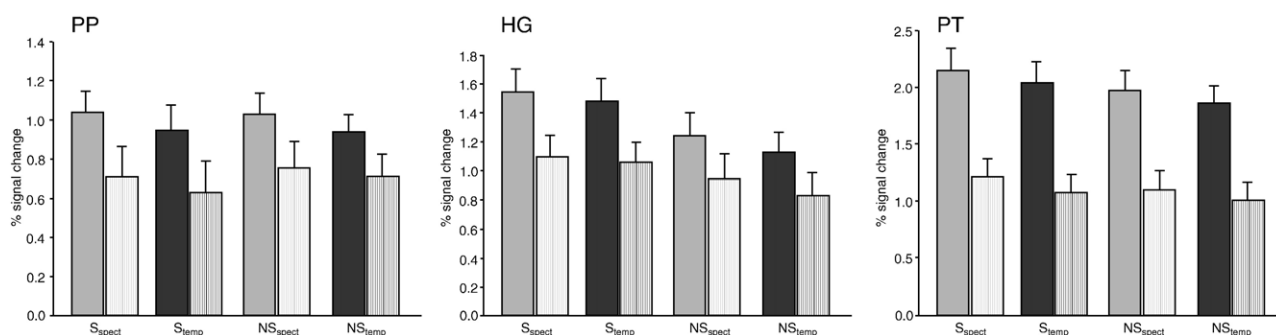

Fig. 2 – Analysis of three bilateral superior temporal plane regions for the main effects of LSC (speech/non-speech), ADC (temporal/spectral) and hemisphere (left/right). Mean beta-values averaged across all subjects within bilateral planum polare (PP), Heschl's gyrus (HG), and planum temporale (PT). Error bars indicate the standard deviation. Light gray bars refer to the temporal conditions; dark gray bars refer to the spectral condition. Bars in plain-color refer to the left; bars in patterned color refer to the right hemisphere.

$p=0.005$), PT ($F(1,15)=22.83$, $p=0.000$) and HG ($F(1,15)=7.15$, $p=0.017$) demonstrating stronger responses in the left as compared to the right hemisphere. For the HG the analysis also revealed a significant main effect of the factor LSC ($F(1,15)=7.47$, $p=0.015$), indicating stronger responses during the speech condition, and a significant main effect of the factor ADC ($F(1,15)=6.17$, $p=0.025$), showing stronger responses for the spectral conditions. No significant interactions could be observed.

2.3.   Main effect of LSC

The speech condition as compared to the non-speech condition activated the auditory cortex in the STS bilaterally, regardless of the acoustic variations. Additionally, frontal activation was observed for the processing of speech sounds in the left middle frontal gyrus (MFG) (cf. Fig. 3 and Table 1). The comparison of the non-speech condition with the speech condition did not reveal significant activation. Fig. 3 also shows the mean percent signal change for two regions of interest (left and right STS). Results of a subsequent 2×2 ANOVA with the factors LSC (speech/non-speech) and ADC (temporal/spectral) separately for each ROI are summarized in Table 2. All regions displayed a

Table 1 – fMRI main effects of factor linguistic stimulus category (LSC)

Condition and anatomical area	L/R	T-value	x	y	z
Speech>Non-speech					
Superior temporal sulcus	L	14.33	−58	−18	−4
	R	9.70	50	−20	−6
Middle frontal gyrus	L	4.60	−40	8	38
Cingulate gyrus	L	7.10	−4	−20	30
	R	4.38	8	−58	38
Thalamus	L	5.83	−12	−14	10
Non-speech>Speech					
No suprathreshold voxels					

The coordinates are given according to the MNI space together with their T-scores.

main effect of the factor LSC, indicating significantly stronger hemodynamic responses to speech than to non-speech stimuli, and a significant effect of the factor ADC, reflecting stronger responses to spectral than temporal variations in both
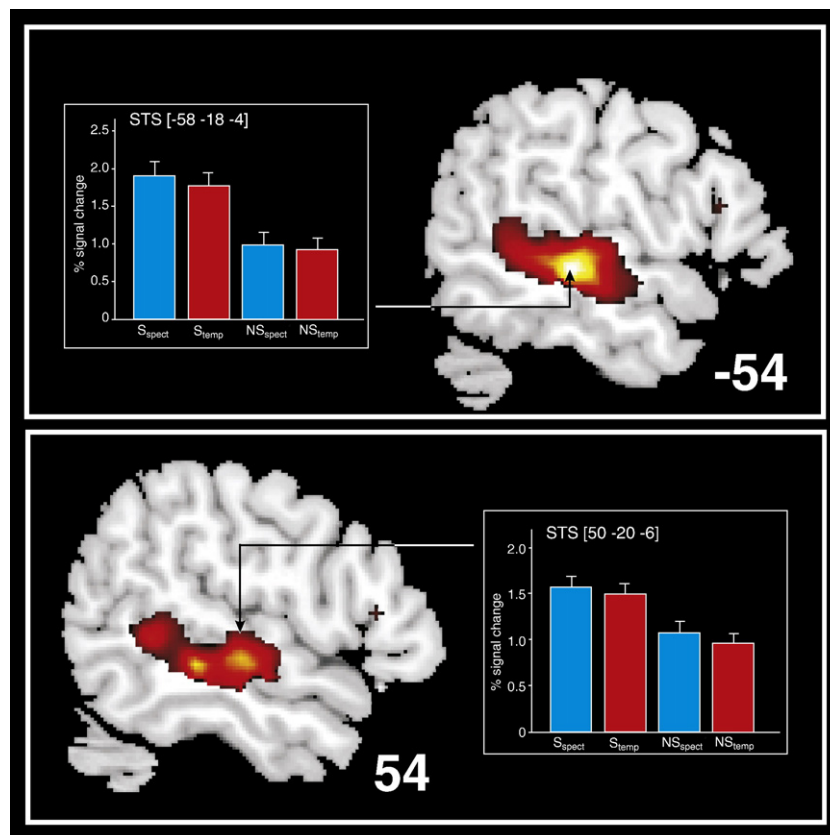

Fig. 3 – Group main effects for the factor LSC: stronger responses to speech sounds are color-coded at a red–yellow scale. Comparison of non-speech sounds in relation to speech sounds did not reveal significant responses. Results are superimposed onto sagittal sections of a standard anatomical template, $x=-54, 54$. Bar plots show mean signal increase for the four conditions, discrimination of non-speech sounds based on temporal features (NS$_{temp}$), discrimination of non-speech sounds based on spectral features (NS$_{spect}$), discrimination of speech sounds based on temporal features (S$_{temp}$), discrimination of speech sounds based on spectral features (S$_{spect}$) in particular regions, right superior temporal sulcus (STS) [50 −20 −6], and left superior temporal sulcus (STS) [−58 −18 −4], exploring stronger responses to speech than to non-speech sounds.

Table 2 – Summary of results of the 2×2 ANOVAs carried out to test the effects of LSC and ADC on the hemodynamic responses in particular regions

ROI	Factor				Interaction	
	LSC	(df: 1.15)	ADC	(df: 1.15)	LSC×ADC	(df:1.15)
	F	p	F	P	F	p
Speech>Non-speech						
STS [−58 −18 −4]	160.41	0.000	5.69	0.031	0.532	0.477
STS [50 −20 −6]	77.02	0.000	10.9	0.006	0.139	0.715
MFG [−40 8 38]	20.26	0.000	3.135	0.096	0.967	0.341
Temporal>Spectral						
IFG [−58 8 14]	0.14	0.9	69.844	0.000	1.563	0.23
PO [−60 −20 25]	0.52	0.48	58.65	0.000	0.861	0.368
FO [−40 22 2]	0.17	0.68	11.31	0.004	9.60	0.007
FO [28 30 −4]	0.004	0.95	30.47	0.000	7.59	0.015
MFG [−44 36 26]	1.68	0.21	13.93	0.002	1.339	0.265
MFG [38 34 20]	0.03	0.86	30.08	0.000	0.065	0.802
Spectral>Temporal						
STS [−62 −24 0]	107.53	0.000	78.62	0.000	0.549	0.47
STS [64 −26 −2]	25.38	0.000	117.08	0.000	0.215	0.649
Anterior STS [−52 −4 −16]	2.63	0.126	23.26	0.000	0.868	0.366
Anterior STS [56 −4 −8]	10.25	0.006	24.19	0.000	1.013	0.33

the speech and non-speech condition. No interaction could be observed.

2.4. Main effect of ADC

As shown in Fig. 4 and listed in Table 3, the temporal condition in comparison to the spectral condition activated the MFG and the frontal operculum (FO) bilaterally, and the parietal operculum (PO) and the IFG of the left hemisphere. Subsequent ROI analyses further corroborate these results (cf. Table 2). All regions displayed a significant main effect of the factor ADC and no main effect of the factor LSC, demonstrating stronger hemodynamic responses during the discrimination task based on temporal features as compared to the discrimination based on spectral features independent of the "speechness" of the sounds. Furthermore, the bilateral FO also displayed a LSC× ACD interaction.

In contrast, comparison of the spectral condition with the temporal condition revealed activations of the anterior and middle parts of the STS, and the middle occipital gyrus of both hemispheres (cf. Fig. 4 and Table 3). As illustrated in Fig. 4 and listed in Table 2, the analysis of ROI data revealed a significant main effect of the factor ADC, demonstrating significantly stronger hemodynamic responses during the categorization based on spectral features as compared to the categorization based on the temporal features in all regions of interest. Additionally, the left middle STS, the right middle STS and the right anterior STS also showed a main effect of the factor LSC, indicating stronger responses to speech than non-speech sounds. No interaction could be observed.

3. Discussion

The present study investigated the functional organization of sublexical auditory perception with respect to the spectro-temporal processing in speech and non-speech sounds. Results show that the discrimination of auditory sounds, when separately compared to a silent baseline, led to significant activation of the bilateral superior temporal plane with a clear functional leftward asymmetry. In former studies we, and others, have shown that areas along the superior temporal plane, namely the HG and PT, are specifically endowed to the processing of rapidly changing acoustic cues in speech and non-speech sounds (Binder et al., 2000; Jancke et al., 2002; Joanisse and Gati, 2003; Meyer et al., 2005; Zaehle et al., 2004). In the present study we confirmed these former results by demonstrating consistent left hemispheric functional asymmetry for all conditions regardless of the ADC or the "speechness" of the sounds. Because all of the stimuli contained rapidly changing spectral information, the data imply that the left superior temporal plane is involved in the fine grained analysis of dynamic auditory information in general and thus appears to be independent from elaborate discrimination tasks. However, since the left hemisphere exhibited a functional dominance for the entire set of stimuli it is not possible to exactly describe which aspect of the auditory signals specifically drives the observed functional asymmetry.

3.1. Effects of LSC

When comparing the processing of speech and non-speech sounds stronger responses were found bilaterally in the middle portion of the superior temporal region with a dominance of the left hemisphere during the processing of speech sounds (cf. Fig. 3). The middle portion of the STS and adjacent areas in the superior temporal gyrus (STG) bilaterally have consistently been implicated in the phonological analysis of speech sounds (Binder et al., 2000; Indefrey and Levelt, 2004; Liebenthal et al., 2005; Mummery et al., 1999; Price et al., 1996; Uppenkamp et al., 2006; Zatorre et al., 1992). As suggested by Hickok and Poeppel (2007), portions of the bilateral STS are driven by acoustic

signals that contain phonological information. Our results are in agreement with these former findings and provide further evidence for the eminent role of the bilateral STS in phonological-level processes during speech perception.

3.2. Effects of ADC

The comparison of the tasks in which participants had to discriminate non-speech and speech stimuli by changes in spectral properties in relation to changes in temporal modifications showed stronger activations of the bilateral STS for spectral based discrimination (cf. Fig. 4). To accomplish this task, subjects had to detect and compare differences in the spectral characteristics of the auditory stimuli. There is a growing body of evidence showing that the processing of spectral properties, being characteristic features of auditory sounds, is associated with activations of the STS in the two hemispheres (Griffiths et al., 1998; Hall et al., 2002; Warren et al., 2006). Furthermore, bilateral STS/STG activations have been reported for the processing of spectrally relevant acoustic properties compared to temporally relevant acoustic properties in non-speech sounds using positron emission tomography (PET) (Zatorre and Belin, 2001) and fMRI (Jamison et al., 2006). Our results are in line with these findings and further support the notion of an anterolaterally directed processing stream specialized for the processing of auditory sound features (Rauschecker, 1998a; Rauschecker and Tian, 2000).

When participants had to discriminate the non-speech and speech stimuli on the basis of temporal variations, a task that requires a segmental analysis of the auditory stimuli, we identified a network comprising the FO and MFG bilaterally, and the PO and the posterior part of the IFG of the left hemisphere. In general, the left frontal cortex plays an important role in phonological processing. As has been shown, the categorical perception of VOT is impaired in non-fluent aphasic patients with left frontal cortex lesions (Blumstein et al., 1977a,b). Japanese-speaking individuals with lesions of the left IFG show impairments in processing phonologically coded Kana while having no difficulties processing orthographically coded Kanji (Sasanuma and Monoi, 1975). Furthermore, several neuroimaging studies demonstrated an involvement of the inferior and middle frontal gyri in the processing of phonemes (Burton et al., 2000; Joanisse and Gati, 2003) and sounds with fast acoustic transients (Joanisse and Gati, 2003; Johnsrude et al., 1997; Poldrack et al., 2001), suggesting specific sensitivity of the left IFG for the processing of speech and non-speech sounds that contain rapidly changing information at a time

range similar to speech. It has been argued that the process of segmenting the initial consonant from the following vowel in speech discrimination may be viewed as the conceptual correlate of inferior frontal activation (Burton, 2001; Burton et al., 2000). The present study confirms the involvement of the left IFG in discriminating speech sounds on the basis of temporal but not spectral feature variations. Thus, the left IFG seems to be dedicated to a segmental analysis of sublexical speech sounds. Furthermore, by obtaining the same pattern of left IFG activation during the discrimination of non-speech sounds on the basis of temporal but not spectral features we provide evidence that the left IFG functioning is not associated with the linguistic nature of the processed auditory sounds.

A further important observation of the present study is the specific activation of an area in the convex of the anterior part of the supramarginal gyrus and the inferior postcentral gyrus constituting the parietal operculum (PO) for the discrimination of speech and non-speech sounds based on temporal features. It has been demonstrated cytoarchitectonically that the PO contains auditory-related cortex and plays a role in auditory function (Galaburda and Sanides, 1980). Furthermore, electrophysiological measurements of the exposed cortex revealed that the PO shows auditory evoked potentials similar to that of the secondary auditory area of the lower lip of the Sylvian fissure (Celesia, 1976). The PO also displays a left–right asymmetry in the surface pattern (Steinmetz et al., 1990) and is larger in the left compared to the right hemisphere in most right-handed subjects (Habib et al., 1995). Morphological deviations of the left PO have been associated with developmental dyslexia (Habib, 2000; Robichon et al., 2000; Silani et al., 2005). Furthermore, a recent fMRI investigation showed activation of the left inferior parietal cortex specifically linked to phonological processes in controls (Dufor et al., 2007). This study also showed that individuals with developmental dyslexia fail to specifically recruit the left PO during phoneme categorization. Additionally, lesion studies identified the left PO as a principal site of phonemic processing in speech perception (Caplan et al., 1995). In the present study we demonstrate that the left PO is involved in the discrimination of speech and non-speech sounds based on temporal features but not by means of the spectral features of the sounds and thus extends the previous view in that it associates the PO with elemental auditory functions.

In sum, the results underscore a particular preference of the left dorsal stream, comprising the posterior portion of the inferior frontal cortex and the parietal operculum, for the processing of temporal features available in the acoustic signal.

Fig. 4 – Group main effects for the factor ADC: stronger responses to temporal processing are color-coded at a red–yellow scale. Starker responses to spectral processing are color-coded at a blue-petrol scale. Results are superimposed onto sagittal sections of a standard anatomical template, $x = -55, -41, 43, 54$. Bar plots show mean signal increase for the four conditions, discrimination of non-speech sounds based on temporal features (NS_{temp}), discrimination of non-speech sounds based on spectral features (NS_{spect}), discrimination of speech sounds based on temporal features (S_{temp}), discrimination of speech sounds based on spectral features (S_{spect}) in particular regions, right middle frontal gyrus (MFG) [38 34 20], left middle frontal gyrus (MFG) [−44 36 26], right frontal operculum (FO) [28 30 −4], left frontal operculum (FO) [−40 22 2], left inferior frontal gyrus (IFG) [−58 8 14], and left parietal operculum (PO) [−60 −20 25] exploring starker responses for the temporal then spectral modulations, and right anterior superior temporal sulcus (STS) [56 −4 −8], left anterior superior temporal sulcus (STS) [−52 −4 −16], right superior temporal sulcus (STS) [64 −26 −2], left superior temporal sulcus (STS) [−62 −24 0] exploring a stronger response to spectral modulations.
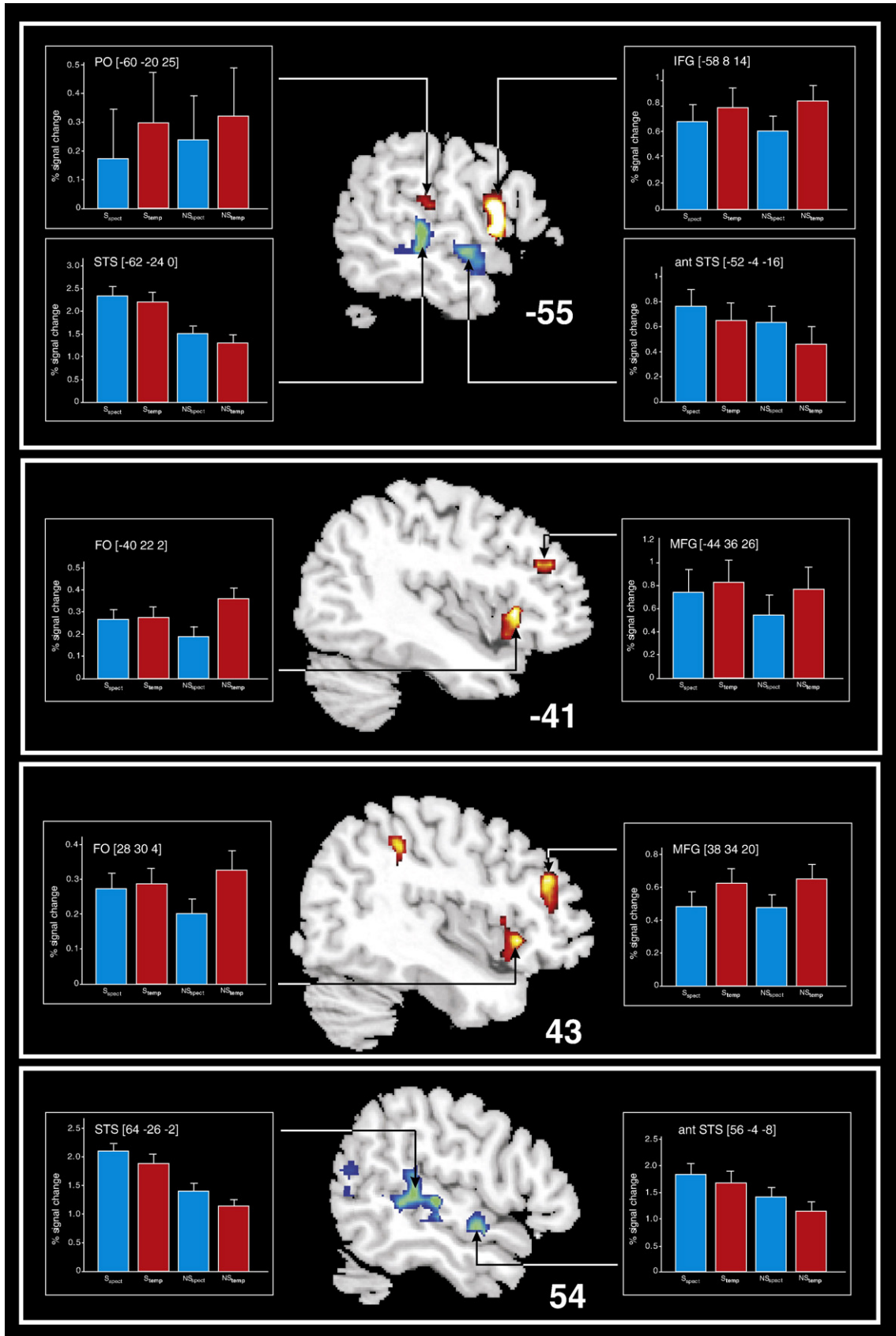

Table 3 – fMRI main effect of factor acoustic discrimination criterion (ADC)

Condition and anatomical area	L/R	T-value	x	y	z
Temporal > Spectral					
Inferior frontal gyrus	L	7.25	−58	8	14
Parietal operculum	L	4.89	−60	−20	25
Middle frontal gyrus	L	4.37	−44	36	26
	R	6.06	38	34	20
	R	5.44	34	44	8
Frontal operculum	L	5.62	−40	22	2
	R	6.62	28	30	−4
Nucleus caudatus	R	5.87	10	4	18
Inferior parietal lobe	R	5.01	40	−38	42
Spectral > Temporal					
Superior temporal sulcus	L	9.35	−62	−24	0
	R	9.11	64	−26	−2
Anterior superior temporal sulcus	L	5.25	−52	−4	−16
	R	5.33	56	−4	−8
Middle occipital gyrus	L	4.25	−40	−78	20
	R	6.17	46	−72	20
Cuneus	L	4.42	−12	−80	40

The coordinates are given according to the MNI space together with their T-scores.

Recent brain imaging studies have provided corroborating evidence for a hemispheric specialization within the auditory cortex by demonstrating preference of left auditory cortical areas for rapidly changing acoustic cues available in speech and non-speech sounds (Jancke et al., 2002; Zaehle et al., 2004) and a dominance of right perisylvian cortex for the processing of spectral profiles (Boemio et al., 2005; Schonwiesner et al., 2005). Our data are in line with these observations in that we revealed a dominance of the left hemispheric auditory cortex in the initial acoustic analysis of speech as well as non-speech sounds with analogous spectro-temporal features. As proposed by Hickok and Poeppel (2007), the auditory signal passes through several stages of analysis. The earliest stage carries out a spectro-temporal analysis and appears in the bilateral, but left-dominant supratemporal plane. The present study confirms this assumption and further evidences the importance of the left HG and PT in the acoustic analysis of auditory signals independent of the linguistic content. After the initial spectro-temporal analysis, the processing system subsequently diverges into a ventral pathway and a dorsal pathway. The ventral pathway is presumed to map sensory or phonological representations onto lexical conceptual representations. In the present study we showed a specific involvement of the middle STS for the processing of consonant–vowel syllables and thus further evidence the elevated role of the bilateral STS in phonological-level processes during speech perception. The dorsal processing stream is assumed to be specifically involved in speech perception tasks like syllable discrimination. Our results demonstrate an involvement of a dorsal processing stream, in particular of the left posterior IFG and left PO during the discrimination of speech and non-speech sounds by means of temporal features. Thus, we suggest that the functioning of the dorsal stream in auditory speech processing is dedicated to auditory segmentation and is independent from the linguistic content of the processed auditory input.

It has been suggested that the dorsal fronto-parietal network, predominantly in the left hemisphere, might be tuned in ontogenesis, because attention to phonetic segments provides the basis for the process of learning speech articulation and is functioning to interface auditory and articulatory representations of speech (Hickok and Poeppel, 2000, 2004, 2007). This is a function that remains intact in adults, providing the basis for acquiring new vocabulary. Interestingly, several studies demonstrated malfunctioning of the left IFG (Temple et al., 2000) and the left PO (Dufor et al., 2007) in children and adults with developmental dyslexia during phonological processing (Cao et al., 2006; Hoeft et al., 2007). It has also been reported that the occurrence of morphological deviations in dyslexic adults in left prefrontal and left parietal areas correlates highly with deficits in processing rapidly presented sounds (Jernigan et al., 1991). Thus, malfunctioning of the left dorsal processing stream during ontogenesis might be a general neurofunctional base for the development of language-related impairments such as dyslexia.

In sum, our data show that spectro-temporal analysis of speech and non-speech sounds is performed at an initial stage by bilateral, but left-dominant superior temporal cortex. Furthermore, we demonstrate that the processing of temporal information in both speech and non-speech sounds that might be related to the extraction of segmental information, specifically engaged a left hemispheric fronto-parietal cortical network. Thus, our fMRI data provide additional evidence for the existence of a dorsal auditory processing stream in speech processing. Moreover, by evidencing that this cortical network is activated when the extraction of temporal or segmental information is required to solve an auditory discrimination task, we found new evidence to enhance the present knowledge of dorsal auditory pathway functions and demonstrate that this pathway is specifically dedicated to serving general segmental acoustic analyses of sounds rather than displaying specificity for speech processing.

4.   Experimental procedure

4.1.   Subjects

Sixteen native speakers of Swiss–German (aged 22–36, mean: 28.43) participated in this study. After a full explanation of the nature and risks of the study, subjects gave informed consent for the participation according to a protocol approved by the local Ethics Committee. They had no history of any neurological, psychiatric or hearing impairment. All subjects were consistently right-handed according to standard handedness questionnaires (Annett, 1992; Steingrüber and Lienert, 1976). In order to avoid gender effects in brain size and shape as well as in functional language lateralization (Luders et al., 2002), only male subjects were included in the study. All subjects were made familiar with the task prior to scanning.

4.2.   Stimulation

Four sets of auditory stimuli were generated (sampling depth: 16 bits, sampling rate of 44.1 kHz) using SoundForge 4.5 Software (Sonic Foundry Inc., http://www.sonicfoundry.com) and PRAAT (http://www.fon.hum.uva.nl/praat/). Two of these four sets of
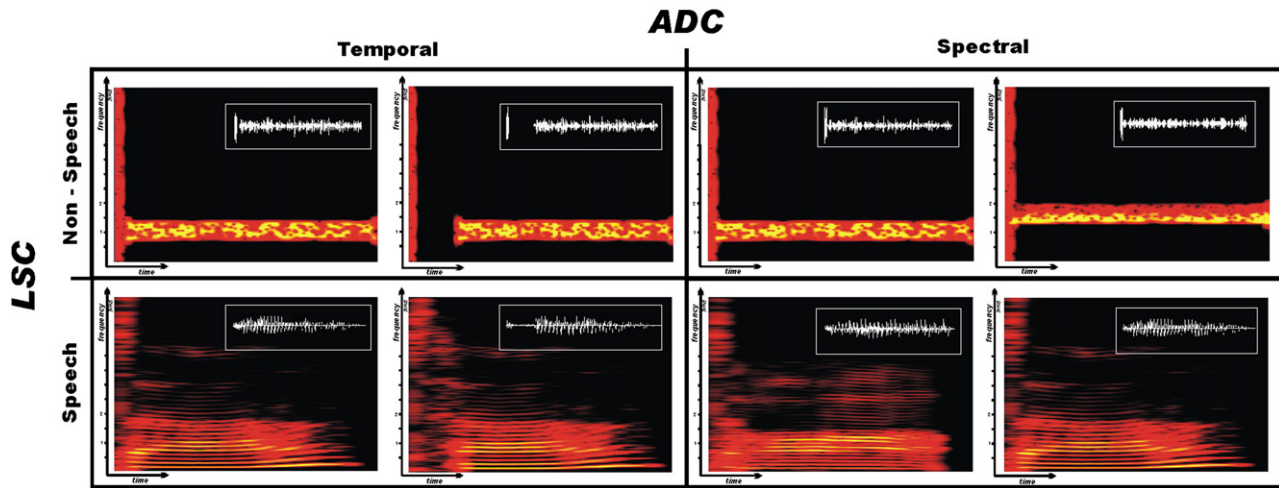
Fig. 5 – Stimulus types used in the four conditions of this experiment: Figure shows spectrograms and waveforms of examples of stimuli pairs illustrating the 2×2 design with the factors linguistic stimulus category (LSC) and acoustic discrimination criterion (ADC).

stimuli were designed to tap the processing of temporal and spectral information without phonological cues (non-speech condition), and two sets with phonological cues (speech condition). For the non-speech condition we used a modified version of the stimulus material used in the study by Zaehle et al. (2004). In particular we created gap stimuli composed of two sound elements separated by a gap. The leading element was a wideband noise burst with a length of 7 ms. The trailing element was a bandpassed noise (width of 500 Hz) with a duration of 300 ms. All noise bursts were shaped with 0.5-ms linear rise–fall times. Temporal information was controlled by varying gap durations, resulting in different noise onset times (NOT) (5, 60, 80, 100 ms). Spectral information was controlled by varying the center frequencies of the trailing element (1000, 1500, 1600, 1700 Hz) with a bandwidth of ±250 Hz. In the former investigation we used stimuli with gap durations of 8 and 32 ms and a fixed center frequencies of the trailing element, while we varied the length of the leading elements in the context of a delayed matching to sample gap detection task (Zaehle et al., 2004). In the speech condition we used recorded consonant–vowel (CV) syllables, spoken by a trained phonetician. To modify the temporal information, different voice onset times (VOT) were used resulting in the perception of different consonants (VOTs in ms/da/=05, /ta/=60). All participants were able to perceive the stimuli as the CV-syllable /ta/ and /da/ as tested by debriefing after initial training session and after the experiment. Variations in

spectral information were achieved by modifying the vowel formants of the CV syllables such that they could be perceived as male or female voices (male /a/: F0: 107 Hz, F1: 851 Hz, F2: 1152 Hz, F3: 2819 Hz; female /a/: F0: 180 Hz, F1: 849 Hz, F2: 1422 Hz, F3: 2762 Hz). The duration of the syllables ranged from 257 to 269 ms. Fig. 5 shows spectrograms and waveforms of example stimulus pairs.

4.3. Experimental design

While lying in the fMRI scanner, participants performed a same–different discrimination task on pairs of speech and non-speech stimuli. Participants had to base their discrimination on temporal or spectral features of the sounds and respond by pressing one of two buttons. Within one trial, participants heard and discriminated five stimulus pairs, all belonging to one of the four experimental conditions (speech or non-speech, temporal or spectral). Fig. 6 shows the timing of an individual trial. For each of the four experimental conditions, the discrimination of non-speech sounds based on temporal (NS_{temp}) and on spectral features (NS_{spect}), and the discrimination of speech sounds based on temporal (S_{temp}) and on spectral features (S_{spect}), 20 trials were presented in randomized order separated into two runs. In particular, we applied for the NS_{temp} condition, 25 "same" pairs (NOT of 5–5) and 75 "different" pairs (NOT of 5–60, 5–80 and 5–100 ms), and in the NS_{spect} condition,
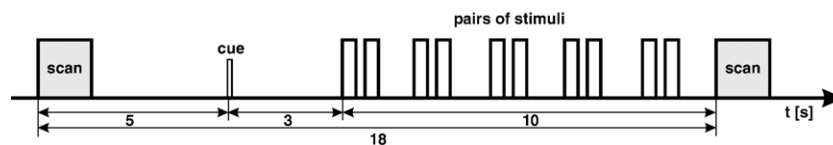

Fig. 6 – Schematic description of behavioral task: the stimulation started with a visual cue, a colored cross which signaled the condition followed 3 s later by the first pair of stimuli consisting of two sounds separated by 500-ms silence. The second pair of stimuli followed 2 s after onset of the first pair and so on. In sum, 5 pairs of stimuli of the same condition were presented within one trial. The task was to compare the two stimuli of one pair and press one of two response keys to indicate that the sounds were the same or different immediately after the presentation of each pair. Two seconds after presentation of the last stimulus pair the acquisition of the one fMRI image started.

25 "same" pairs (bandpass of 1–1 kHz) and 75 "different" pairs (bandpass of 1–1.5, 1–1.6 and 1–1.7 kHz). Similarly in the S_{temp} condition, 24 "same" pairs (VOT of 5–5, 60–60) and 76 "different" pairs (VOT of 5–60), and in the S_{spect} condition, 24 "same" pairs (male–male, female–female) and 76 "different" pairs (male–female) were applied. All stimuli were presented in a randomized and counterbalanced order. Twenty-one trials of a resting silence condition were included as a baseline in each run. During scanning, participants viewed a fixation cross on a projection screen with the aid of an angled mirror placed above their eyes. Binaural auditory stimulation was achieved by a digital playback system including a high-frequency shielded transducer system.

4.4. Data acquisition

Images were acquired using a 3-T whole body scanner (Philips Intera) equipped with an eight-channel Philips head coil, radiofrequency transmission and signal reception. Using a mid-sagittal scout image, 16 axial slices (slice thickness 5 mm) of a T2*-weighed gradient echo EPI sequence were acquired parallel to the bicommissural plane and covering the whole brain. The following acquisition parameters were used: repetition time (TR): 18 s, echo time (TE): 35 ms, flip angle: 90°, voxel size: $1.72 \times 1.72 \times 5$ mm^3. Data acquisition was clustered in the first 1.5 s of the TR, leaving 16.5 s for auditory stimulation without gradient noise (Hall et al., 1999) providing the advantage of silent data acquisition in auditory fMRI designs (Gaab et al., 2007a,b; Schmidt et al., 2007; Zaehle et al., 2007). A total of 122 volumes divided into two runs each lasting 18.3 min were acquired in the course of the experiment. Additionally, a high-resolution anatomical scan was acquired for each participant.

4.5. Data analysis

Artifact elimination and image analysis was performed using MATLAB 7 (Mathworks Inc., Natiek, MA, USA), the SPM99 software package (Institute of Neurology, London, UK), and the MarsBar toolbox for SPM (http://marsbar.sourceforge.net/) (Brett et al., 2002). Preprocessing of functional MR images included movement correction and coregistration with the anatomical data. The scans were then normalized to a standardized brain space (defined by the Montreal Neurological Institute) using a non-linear transformation matrix (voxel size $2 \times 2 \times 2$ mm) and smoothed using a Gaussian kernel with 6-mm full-width at half-maximum (FWHM). The EPI time series was analyzed using a general linear model analysis (Friston et al., 1995). Separate regressors were defined for each condition using a box-car function without convolving with the canonical hemodynamic response function as reference waveform. After estimation of the model parameters for each subject, an analysis of variance was calculated for the whole group, using the individual contrast images for the main effects. In a first step, we compared the four conditions separately with the resting baseline. The results of these statistics were thresholded by $T = 7.7$ ($p = 0.05$ corrected for multiple comparisons). Region-of-interest (ROI) analyses were subsequently conducted in order to quantify the functional hemisphere differences of local brain activity in

auditory fields stretching along the supratemporal plane (STP) as a function of auditory processing. For that purpose a set of six distinct ROIs for each subject was placed in HG (middle STP), PP (anterior STP) and PT (posterior STP) of each hemisphere. ROIs were anatomically defined based on macroanatomical landmarks (Rademacher et al., 2001; Steinmetz et al., 1990). Mean beta-values were collected for each subject from each ROI (HG, PP, PT) and subjected to a repeated-measure within-subjects ($2 \times 2 \times 2$) ANOVA with the factors linguistic stimulus category (LSC) (speech/non-speech), acoustic discrimination criterion (ADC) (temporal/spectral), and hemisphere (left/right).

In a next step subtractions between conditions were performed. The two principle results discussed are those that evaluate the statistical significance of the two main effects. The effect of ADC was analyzed by comparing the processing of both speech and non-speech sounds based on temporal variations with the processing of these sounds based on spectral variations. The effect of LSC was evaluated by comparing the processing of the speech sounds with the processing of the non-speech sounds regardless of the temporal or spectral variations. Results of these statistics were thresholded by $T = 3.73$ ($p = 0.001$ uncorrected for multiple comparisons) and $k = 30$ voxels. Additionally, to further explore the statistical results of these main effects, functionally defined ROIs were examined. Here we used the peak-activated cortical voxel as the center and a sphere with a radius of 6 mm. To test for main effects of ADC and LSC, 2×2 ANOVAs with the factors LSC (speech/non-speech) and ADC (temporal/spectral) were run for each of these ROIs.

R E F E R E N C E S

Annett, M., 1992. Five tests of hand skill. Cortex 28, 583–600.
Binder, J.R., Frost, J.A., Hammeke, T.A., Cox, R.W., Rao, S.M., Prieto, T., 1997. Human brain language areas identified by functional magnetic resonance imaging. J. Neurosci. 17, 353–362.
Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S., Springer, J.A., Kaufman, J.N., Possing, E.T., 2000. Human temporal lobe activation by speech and nonspeech sounds. Cereb. Cortex 10, 512–528.
Blumstein, S.E., Baker, E., Goodglass, H., 1977a. Phonological factors in auditory comprehension in aphasia. Neuropsychologia 15, 19–30.
Blumstein, S.E., Cooper, W.E., Zurif, E.G., Caramazza, A., 1977b. The perception and production of voice-onset time in aphasia. Neuropsychologia 15, 371–383.
Boemio, A., Fromm, S., Braun, A., Poeppel, D., 2005. Hierarchical and asymmetric temporal sensitivity in human auditory cortices. Nat. Neurosci. 8, 389–395.
Brett, M., Anton, J.-L., Valabregue, R., Poline, J.-B., 2002. Region of Interest Analysis Using an SPM Toolbox, 16 ed.
Burton, M., 2001. The role of inferior frontal cortex in phonological processing. Cogn. Sci. 25, 695–709.

Burton, M.W., Small, S.L., 2006. Functional neuroanatomy of segmenting speech and nonspeech. Cortex 42, 644–651.

Burton, M.W., Small, S.L., Blumstein, S.E., 2000. The role of segmentation in phonological processing: an fMRI investigation. J. Cogn. Neurosci. 12, 679–690.

Cao, F., Bitan, T., Chou, T.L., Burman, D.D., Booth, J.R., 2006. Deficient orthographic and phonological representations in children with dyslexia revealed by brain activation patterns. J. Child Psychol. Psychiatry 47, 1041–1050.

Caplan, D., Gow, D., Makris, N., 1995. Analysis of lesions by MRI in stroke patients with acoustic–phonetic processing deficits. Neurology 45, 293–298.

Celesia, G.G., 1976. Organization of auditory cortical areas in man. Brain 99, 403–414.

Crinion, J.T., Lambon-Ralph, M.A., Warburton, E.A., Howard, D., Wise, R.J., 2003. Temporal lobe regions engaged during normal speech comprehension. Brain 126, 1193–1201.

Dufor, O., Serniclaes, W., Sprenger-Charolles, L., Demonet, J.F., 2007. Top–down processes during auditory phoneme categorization in dyslexia: a PET study. NeuroImage 34, 1692–1707.

Efron, R., 1963. Temporal perception, aphasia and D'ej'a vu. Brain 86, 403–424.

Fiez, J.A., Raichle, M.E., Balota, D.A., Tallal, P., Petersen, S.E., 1996. PET activation of posterior temporal regions during auditory word presentation and verb generation. Cereb. Cortex 6, 1–10.

Friston, K.J., Holmes, A.P., Poline, J.B., Grasby, P.J., Williams, S.C., Frackowiak, R.S., Turner, R., 1995. Analysis of fMRI time-series revisited. NeuroImage 2, 45–53.

Gaab, N., Gabrieli, J.D., Glover, G.H., 2007a. Assessing the influence of scanner background noise on auditory processing: I. An fMRI study comparing three experimental designs with varying degrees of scanner noise. Hum. Brain Mapp. 28, 703–720.

Gaab, N., Gabrieli, J.D., Glover, G.H., 2007b. Assessing the influence of scanner background noise on auditory processing: II. An fMRI study comparing auditory processing in the absence and presence of recorded scanner noise using a sparse design. Hum. Brain Mapp. 28, 721–732.

Galaburda, A., Sanides, F., 1980. Cytoarchitectonic organization of the human auditory cortex. J. Comp. Neurol. 190, 597–610.

Griffiths, T.D., Buchel, C., Frackowiak, R.S., Patterson, R.D., 1998. Analysis of temporal structure in sound by the human brain. Nat. Neurosci. 1, 422–427.

Habib, M., 2000. The neurological basis of developmental dyslexia: an overview and working hypothesis. Brain 123 (Pt 12), 2373–2399.

Habib, M., Robichon, F., Levrier, O., Khalil, R., Salamon, G., 1995. Diverging asymmetries of temporo-parietal cortical areas: a reappraisal of Geschwind/Galaburda theory. Brain Lang. 48, 238–258.

Hall, D.A., Haggard, M.P., Akeroyd, M.A., Palmer, A.R., Summerfield, A.Q., Elliott, M.R., Gurney, E.M., Bowtell, R.W., 1999. "Sparse" temporal sampling in auditory fMRI. Hum. Brain Mapp. 7, 213–223.

Hall, D.A., Johnsrude, I.S., Haggard, M.P., Palmer, A.R., Akeroyd, M.A., Summerfield, A.Q., 2002. Spectral and temporal processing in human auditory cortex. Cereb. Cortex 12, 140–149.

Hickok, G., Poeppel, D., 2000. Towards a functional neuroanatomy of speech perception. Trends Cogn. Sci. 4, 131–138.

Hickok, G., Poeppel, D., 2004. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition 92, 67–99.

Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nat. Rev., Neurosci. 8, 393–402.

Hoeft, F., Meyler, A., Hernandez, A., Juel, C., Taylor-Hill, H., Martindale, J.L., McMillon, G., Kolchugina, G., Black, J.M., Faizi, A., Deutsch, G.K., Siok, W.T., Reiss, A.L., Whitfield-Gabrieli, S., Gabrieli, J.D., 2007. Functional and morphometric brain dissociation between dyslexia and reading ability. Proc. Natl. Acad. Sci. U. S. A. 104, 4234–4239.

Indefrey, P., Levelt, W.J., 2004. The spatial and temporal signatures of word production components. Cognition 92, 101–144.

Jamison, H.L., Watkins, K.E., Bishop, D.V., Matthews, P.M., 2006. Hemispheric specialization for processing auditory nonspeech stimuli. Cereb. Cortex 16, 1266–1275.

Jancke, L., Wustenberg, T., Scheich, H., Heinze, H.J., 2002. Phonetic perception and the temporal cortex. NeuroImage 15, 733–746.

Jernigan, T.L., Hesselink, J.R., Sowell, E., Tallal, P.A., 1991. Cerebral structure on magnetic resonance imaging in language- and learning-impaired children. Arch. Neurol. 48, 539–545.

Joanisse, M.F., Gati, J.S., 2003. Overlapping neural regions for processing rapid temporal cues in speech and nonspeech signals. NeuroImage 19, 64–79.

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.

Liebenthal, E., Binder, J.R., Spitzer, S.M., Possing, E.T., Medler, D.A., 2005. Neural substrates of phonemic perception. Cereb. Cortex 15, 1621–1631.

Luders, E., Steinmetz, H., Jancke, L., 2002. Brain size and grey matter volume in the healthy human brain. NeuroReport 13, 2371–2374.

Meyer, M., Zaehle, T., Gountouna, V.E., Barron, A., Jancke, L., Turk, A., 2005. Spectro-temporal processing during speech perception involves left posterior auditory cortex. NeuroReport 16, 1985–1989.

Mummery, C.J., Ashburner, J., Scott, S.K., Wise, R.J., 1999. Functional neuroimaging of speech perception in six normal and two aphasic subjects. J. Acoust. Soc. Am. 106, 449–457.

Poeppel, D., 2001. Pure word deafness and the bilateral processing of the speech code. Cogn. Sci. 25, 679–693.

Poeppel, D., 2003. The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. Speech Commun. 41, 245–255.

Poldrack, R.A., Temple, E., Protopapas, A., Nagarajan, S., Tallal, P., Merzenich, M., Gabrieli, J.D., 2001. Relations between the neural bases of dynamic auditory processing and phonological processing: evidence from fMRI. J. Cogn. Neurosci. 13, 687–697.

Price, C.J., Wise, R.J., Warburton, E.A., Moore, C.J., Howard, D., Patterson, K., Frackowiak, R.S., Friston, K.J., 1996. Hearing and saying. The functional neuro-anatomy of auditory word processing. Brain 119 (Pt 3), 919–931.

Rademacher, J., Morosan, P., Schormann, T., Schleicher, A., Werner, C., Freund, H.J., Zilles, K., 2001. Probabilistic mapping and volume measurement of human primary auditory cortex. NeuroImage 13, 669–683.

Rauschecker, J.P., 1998a. Cortical processing of complex sounds. Curr. Opin. Neurobiol. 8, 516–521.

Rauschecker, J.P., 1998b. Parallel processing in the auditory cortex of primates. Audiol. Neuro-otol. 3, 86–103.

Rauschecker, J.P., Tian, B., 2000. Mechanisms and streams for processing of "what" and "where" in auditory cortex. Proc. Natl. Acad. Sci. U. S. A. 97, 11800–11806.

Robichon, F., Levrier, O., Farnarier, P., Habib, M., 2000. Developmental dyslexia: atypical cortical asymmetries and functional significance. Eur. J. Neurol. 7, 35–46.

Sasanuma, S., Monoi, H., 1975. The syndrome of Gogi (word meaning) aphasia. Selective impairment of kanji processing. Neurology 25, 627–632.

Schmidt, C.F., Zaehle, T., Meyer, M., Geiser, E., Boesiger, P., Jancke, L., 2007. Silent and continuous fMRI scanning differentially modulate activation in an auditory language comprehension task. Hum. Brain Mapp. 29 (1), 46–56 (doi:10.1002/hbm.20372).

Schonwiesner, M., Rubsamen, R., von Cramon, D.Y., 2005. Hemispheric asymmetry for spectral and temporal processing

in the human antero-lateral auditory belt cortex. Eur. J. Neurosci. 22, 1521–1528.

Scott, S.K., Blank, C.C., Rosen, S., Wise, R.J., 2000. Identification of a pathway for intelligible speech in the left temporal lobe. Brain 123 (Pt 12), 2400–2406.

Scott, S.K., Wise, R.J., 2004. The functional neuroanatomy of prelexical processing in speech perception. Cognition 92, 13–45.

Silani, G., Frith, U., Demonet, J.F., Fazio, F., Perani, D., Price, C., Frith, C.D., Paulesu, E., 2005. Brain abnormalities underlying altered activation in dyslexia: a voxel based morphometry study. Brain 128, 2453–2461.

Steingrüber, H.J., Lienert, G.A., 1976. Hand-Dominanz-Test (HDT). Hogrefe, Göttingen.

Steinmetz, H., Ebeling, U., Huang, Y.X., Kahn, T., 1990. Sulcus topography of the parietal opercular region: an anatomic and MR study. Brain Lang. 38, 515–533.

Tallal, P., Piercy, M., 1973. Defects of non-verbal auditory perception in children with developmental aphasia. Nature 241, 468–469.

Temple, E., Poldrack, R.A., Protopapas, A., Nagarajan, S., Salz, T., Tallal, P., Merzenich, M.M., Gabrieli, J.D., 2000. Disruption of the neural response to rapid acoustic stimuli in dyslexia: evidence from functional MRI. Proc. Natl. Acad. Sci. U. S. A. 97, 13907–13912.

Tervaniemi, M., Medvedev, S.V., Alho, K., Pakhomov, S.V., Roudas, M.S., van Zuijen, T.L., Naatanen, R., 2000. Lateralized automatic auditory processing of phonetic versus musical information: a PET study. Hum. Brain Mapp. 10, 74–79.

Uppenkamp, S., Johnsrude, I.S., Norris, D., Marslen-Wilson, W., Patterson, R.D., 2006. Locating the initial stages of speech–sound processing in human temporal cortex. NeuroImage 31, 1284–1296.

Warren, J.D., Scott, S.K., Price, C.J., Griffiths, T.D., 2006. Human brain mechanisms for the early analysis of voices. NeuroImage 31, 1389–1397.

Wernicke, C., 1874. Der aphasische Symptomkomplex: eine psychologische Studie auf anatomischer basis. The Dual Brain: Hemispheric Specialization in Humans. Cohn & Weigert, Breslau.

Zaehle, T., Wustenberg, T., Meyer, M., Jancke, L., 2004. Evidence for rapid auditory perception as the foundation of speech processing: a sparse temporal sampling fMRI study. Eur. J. Neurosci. 20, 2447–2456.

Zaehle, T., Schmidt, C.F., Meyer, M., Baumann, S., Baltes, C., Boesiger, P., Jancke, L., 2007. Comparison of "silent" clustered and sparse temporal fMRI acquisitions in tonal and speech perception tasks. NeuroImage 37, 1195–1204.

Zatorre, R.J., Belin, P., 2001. Spectral and temporal processing in human auditory cortex. Cereb. Cortex 11, 946–953.

Zatorre, R.J., Evans, A.C., Meyer, E., Gjedde, A., 1992. Lateralization of phonetic and pitch discrimination in speech processing. Science 256, 846–849.

Zatorre, R.J., Belin, P., Penhune, V.B., 2002. Structure and function of auditory cortex: music and speech. Trends Cogn. Sci. 6, 37–46.