Spectral and Temporal Processing in Human Auditory Cortex

Deborah A. Hall, Ingrid S. Johnsrude[1,2], Mark P. Haggard, Alan R. Palmer, Michael A. Akeroyd and A. Quentin Summerfield

MRC Institute of Hearing Research, University Park, Nottingham NG7 2RD, [1]MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 2EF and [2]The Functional Imaging Laboratory, Wellcome Department of Cognitive Neurology, St Johns House, 12 Queen Square, London WC1N 3BG, UK

Hierarchical processing suggests that spectrally and temporally complex stimuli will evoke more activation than do simple stimuli, particularly in non-primary auditory fields. This hypothesis was tested using two tones, a single frequency tone and a harmonic tone, that were either static or frequency modulated to create four stimuli. We interpret the location of differences in activation by drawing comparisons between fMRI and human cytoarchitectonic data, reported in the same brain space. Harmonic tones produced more activation than single tones in right Heschl's gyrus (HG) and bilaterally in the lateral supratemporal plane (STP). Activation was also greater to frequency-modulated tones than to static tones in these areas, plus in left HG and bilaterally in an anterolateral part of the STP and the superior temporal sulcus. An elevated response magnitude to both frequency-modulated tones was found in the lateral portion of the primary area, and putatively in three surrounding non-primary regions on the lateral STP (one anterior and two posterior to HG). A focal site on the posterolateral STP showed an especially high response to the frequency-modulated harmonic tone. Our data highlight the involvement of both primary and lateral non-primary auditory regions.

Introduction

Hierarchical acoustical analysis predicts an increasing sensitivity to the abstract constituent features of auditory stimuli as one progresses from primary to non-primary auditory cortical fields. According to this scheme, a greater response to modulated than to static sounds (irrespective of the carrier signal) may be found in discrete higher auditory areas, whereas early areas may respond to primitive aspects of auditory stimuli, such as their frequency spectra. Functional MRI studies, combined with the systematic variation of acoustic stimuli, permit the investigation of functional selectivity in different regions of the human auditory cortex and are of most scientific value when the location of activity is not merely identified by gross morphology, but is also rooted in some understanding of the cortical anatomy as defined by correlation with physiological features. A growing number of anatomical studies have identified discrete auditory sub-regions in humans (Galaburda and Sanides, 1980; Rivier and Clarke, 1997) by the distribution and density of cells within the laminae. These sub-regions appear to be broadly homologous with those observed in primates (Pandya, 1995; Hackett *et al.*, 1998). Moreover, cytoarchitectural borders usually correspond well with physiological borders (Morel *et al.*, 1993), as found in the primary auditory cortex (Merzenich and Brugge, 1973; Liegois-Chauvel *et al.*, 1991). We can use this anatomical knowledge to generate hypotheses about the spatial distribution of functional responses that can be tested in humans non-invasively using fMRI.

A feature of the auditory cortex that is common across species is the presence of bands of frequency-tuned neurons forming a tonotopic gradient within primary and adjacent fields (Scheich 1991; Wallace *et al.*, 1999; Talavage *et al.*, 2000). In cats and primates, neurons in the primary field (A1) show sharper frequency tuning and lower thresholds to pure tone stimulation than do neurons in surrounding non-primary auditory fields (Schreiner and Cynader, 1984; Rauschecker *et al.*, 1995). Consequently, non-primary auditory fields are more strongly activated by sounds, such as harmonic tones and noise, that have broad frequency spectra, than by single frequency tones (Wessinger *et al.*, 2001). Systematic response differences between primary and non-primary auditory fields are also observed to temporal acoustical features, indicating, for example, selectivity for particular modulation rates in non-primary auditory fields. In cats, the non-primary anterior auditory field responds more strongly to higher rates of modulation than does A1 (Tian and Rauschecker, 1994), and the posterior field responds more strongly to slower rates of modulation than does A1 (Heil and Irvine, 1998; Tian and Rauschecker, 1998). These temporal characteristics occur across different shapes of modulation waveform (Schreiner and Urbas, 1988), indicating some degree of independence between temporal and spectral coding. Hierarchical auditory analysis provides a parsimonious explanation for these results (Rauschecker, 1998a,b), in contrast to the alternative view that each class of stimulus activates in its own 'specific' auditory area. In a hierarchy, basic constituent elements of the stimulus provide input to higher auditory fields that extract successively more abstract and complex features; this conceptual framework having its precedent in the visual system (Moutoussis and Zeki, 1997).

A recent hypothesis suggests one hierarchically organized auditory pathway originating in posterolateral areas for sound localization, and another originating in anterolateral areas, involved in the analysis of spectral and temporal features required for identifying acoustical patterns (Rauschecker, 1998a,b; Romanski *et al.*, 1999a; Rauschecker and Tian 2000). The present study measured differential condition-specific responses in primary and non-primary areas in response to acoustical stimuli manipulated for their spectro-temporal pattern to seek evidence for the anterolateral stream in humans.

Materials and Methods

Stimuli

Four classes of acoustical stimuli were defined by crossing two types of carrier signal (single tone and harmonic tone) with two types of modulation (static and frequency modulated, FM). The carrier signal was either a single frequency tone at 500 Hz or a harmonic complex with a fundamental frequency of 186 Hz and components at 186, 372, 558, 744, 930 and 1116 Hz. The spectrum of the harmonic tone was 2.6 octaves wide. There are a number of ways to modulate a carrier signal including frequency, amplitude and phase. The stimuli used here were frequency modulated. The signal, $x(t)$, was either not modulated,

$$x(t) = A_c \sin(2\pi f_c) \quad (1)$$

© Oxford University Press 2002. All rights reserved.

Downloaded from https://academic.oup.com/cercor/article/12/2/140/301066 by guest on 21 August 2022

or frequency modulated cosinusoidally, at a rate f_m of 5 Hz and a depth β of 10:

$$x(t) = A_c\sin(2\pi f_c + \beta\cos(2\pi f_{mt})) \quad (2)$$

where A_c and f_c are the amplitude and frequency, respectively, of the carrier component. The chosen f_m = 5 Hz falls within the distribution of modulation rates found in natural speech. The process of frequency modulation results in additional sideband frequencies centred on each carrier component. For harmonic-complex tones, the valley between harmonics may be filled at large modulation depths, as the upper sidebands of one harmonic overlap with the lower sidebands of the next harmonic. Thus, for the modulated signals, β = 10 was selected in order to preserve the valleys in the frequency spectra (Fig. 1). The difference between a frequency-modulated (FM) and a static tone defined temporal complexity and the difference between a harmonic-complex and a single tone defined spectral complexity. Other acoustical properties were controlled in the stimulus generation, although the harmonic-complex and single tones differed in perceived pitch strength. The internal representation of harmonic-complex tones (Fig. 1*C*,*D*) also included a component of amplitude modulation at a rate of 186 Hz, generated by adjacent higher-frequency harmonics beating together within the same auditory filter to produce a modulated envelope in the neural activation pattern. A difference in amplitude modulation between stimuli was likely to be less detectable by our listeners than a difference in pitch strength. The amplitude of each component, A_c, in the unmodulated signals was adjusted individually during synthesis so that each component was 81 phons, as calculated using a computational model (Moore *et al.*, 1997). The same amplitudes were used for the frequency-modulated stimuli. All the stimuli were of 900 ms duration including 100 ms inverted-cosine ramps at onset and offset.

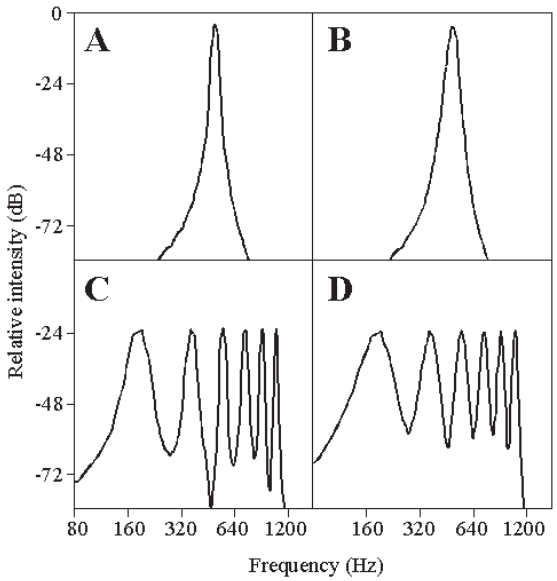

Figure 1. Spectral characteristics of the acoustic stimuli generated using a Fourier analysis of the temporal waveform. (*A*) represents the static single tone and (*B*) the frequency-modulated single tone. (*C*) represents the static harmonic-complex tone and (*D*) the frequency-modulated harmonic-complex tone. Note that the chosen depth of frequency modulation preserved a similar frequency spectrum across static and modulated tones.

Signals were presented diotically over electrostatic headphones that give high-fidelity signal transduction. The transducers were driven by a specially engineered fMRI sound system (Palmer *et al.*, 1998) and were built into industrial ear defenders that provide 10–40 dB of passive attenuation for frequencies between 300 and 6000 Hz. The airborne level of the scanner noise was estimated to be in the region of 88–93 dB SPL at the ear. The presentation levels of both modulated and unmodulated signals were 94 dB SPL (single tone) and 84 dB SPL (harmonic complex) at the ear, calibrated using a KEMAR manikin (Burkhard and Sachs, 1975) equipped with a Brüel and Kjær microphone and measuring amplifier.

Task

Conditions were presented in 16 s blocks. Each stimulus block comprised 15 tone bursts each separated by 100 ms periods of silence. The order of the stimulus conditions was counter-balanced. A silent condition occurred after every four stimulus conditions, during which the baseline level of activation was measured (Fig. 2). The experiment was presented in two 32 min runs in the same session. In a single run, each of the five conditions occurred 24 times, with the order of presentation of the two runs being counter-balanced across subjects.

All listening conditions required target discrimination where the target was a long tone burst (i.e. its duration was 1900 ms, which was equal to the length of two consecutive reference tone bursts plus the intervening 100 ms gap) (Fig. 2). Targets and reference tones were presented in a random sequence, with a ratio of 1:7, so that there was an unpredictable number of targets in each sequence. Subjects were instructed to press a button with a finger on their right hand whenever a target occurred. The sound system logged the occurrence of targets and button responses for off-line analysis.

Imaging Protocol

The study was performed on a 2 T Magnetom VISION (Siemens, Erlangen, Germany) whole-body MRI system, equipped with a head volume coil, at the Functional Imaging Laboratory in Queen Square, London. The subject's head was immobilized using foam pads that applied pressure onto the headphones. Sets of 30 images, providing almost whole-head coverage, were acquired in 2 s using a BOLD-contrast imaging sequence (T_E = 40 ms). The imaging plane was aligned with the Sylvian fissure as identified visually from a sagittal image. Voxel resolution was $3 \times 3 \times 2$ mm with a 0.5 mm inter-slice thickness and the matrix size was 64×64. Acoustical interference, caused by the scanner noise, was reduced by using an inter-scan interval of 8 s, so that any specific auditory response to the scanner noise was likely to have decayed before the subsequent image acquisition (Hall *et al.*, 2000a). A 5 cm wide saturation band was applied coronally across the eyeballs (and frontal poles) to null the MR signal from the eyes which can otherwise produce high variance noise in

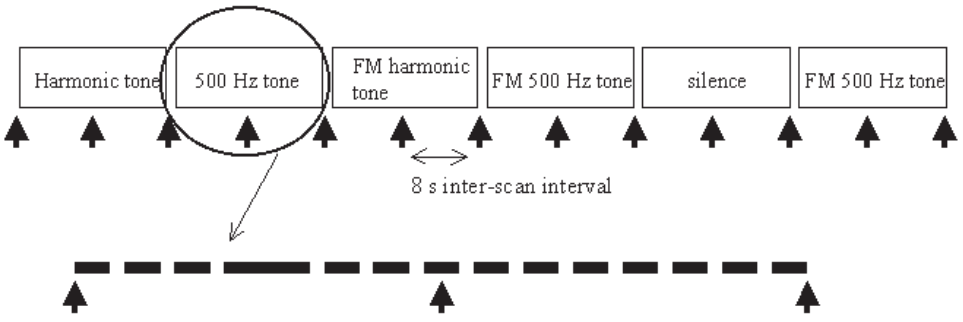

Figure 2. A schematic diagram of the experimental protocol. The upper series represents a sequence of stimulus blocks, with the arrows representing scan acquisitions. Each block comprised a sequence of tone bursts of the same stimulus type with longer duration tones occurring randomly in the sequence. This is illustrated in the lower panel.

Downloaded from https://academic.oup.com/cercor/article/12/2/140/301066 by guest on 21 August 2022

regions of interest. A T_1-weighted, high-resolution brain image (voxel size = $1 \times 1 \times 1.5$ mm) was acquired using an MPRAGE sequence.

Subjects

Data were obtained from six subjects (four male and two female) whose ages ranged from 28 to 49 years. One of the female participants was left-handed (subject 4). None of the subjects had any history of neurological or auditory impairment. All subjects gave informed written consent for participation.

Image Analysis

Images were analysed using SPM99 software (http://www.fil.ion.ucl.ac.uk). The first three volumes in each run were discarded prior to statistical analysis as their purpose was to stabilize the MR signal. For each subject, the remaining images for the two runs were realigned to the first image in the sequence to correct for three-dimensional head movement. Due to the missing brain signal from the frontal cortex, normalization accuracy was enhanced by geometrically aligning the functional and structural images (Maes *et al.*, 1997), including the orientation of the anterior–posterior commissural line. Structural and functional images were then spatially transformed using a T_1-weighted template in a standard brain space used by the International Consortium for Brain Mapping (http://nessus.loni.ucla.edu/icbm) and closely similar to that of Talairach and Tournoux (Talairach and Tournoux, 1988) in the auditory regions. Normalization accuracy was verified by visually comparing each normalized image with the template image to ensure agreement between principle landmarks of the auditory cortex. Furthermore, the close alignment of Heschl's gyrus (HG) across the six subjects was confirmed by the clear delineation of this gyrus in the mean structural image, averaged across the group. Normalized images were re-sampled to a voxel resolution of $2 \times 2 \times 2$ mm and spatially smoothed using a Gaussian kernel of 8 mm full width at half maximum (FWHM). Low-frequency artefacts including aliased respiratory and cardiac effects were removed up to a maximum frequency of 0.38 cycles/minute.

Fixed effects analysis is appropriate for drawing inferences about the typical qualitative aspects of normal functional anatomy within a population sample (Friston *et al.*, 1999). The image data were therefore subjected to a multi-subject, fixed-effects analysis which modelled the stimulus conditions, six realignment parameters and the target-discrimination performance as covariates. Comparisons between conditions were performed to investigate the effects of spectral and temporal complexity and these were specified using combinations of linear contrasts. The two principal results discussed are those that evaluate the statistical significance of the two main effects (the effect of the harmonic and the effect of the FM). To test these effects, we adopted a stringent type of T contrast, a conjunction analysis, that is conceptually similar to performing a Boolean 'AND' operation on two independent T contrasts (Price and Friston, 1997). For the group analysis, each voxel satisfied this criterion with $P < 0.05$, after correction for the performance of multiple T tests within the entire brain space (in this analysis, the brain space included 603 resolution elements, hence correction was applied for this number of independent tests).

For plotting the response matrices, bilateral windows were specified that were parallel to the STP and measured 38×48 mm in X and Y dimensions, thus including primary and non-primary anatomical regions (Fig. 3). For each subject and for each hemisphere, the beta parameters, which reflect a standardized magnitude of each stimulus effect, were extracted for all voxels within these windows. Beta parameters are derived from the general linear model implemented in SPM99 and correspond to the model predictions of the amplitude of condition-specific activation. Beta parameters were then averaged over subjects to yield group-averaged data. Two-dimensional plots of the spatial distribution of the magnitude of the betas were created for both left and right hemispheres. We assume that the standard errors differ little across the region of interest because the data have been spatially smoothed.

Anatomical Identification of Auditory Areas

In humans, architectonic borders of cortical fields do not precisely align with macro-anatomical landmarks and so comparisons between functional activation and anatomy are at best approximate (Rademacher *et al.*, 1993). Here, we do not attempt to use architectonic data to define functional borders, but simply to estimate whether or not activation is likely to occur within a particular region. The principles of this method have previously been applied to evaluate the spatial organization of frequency-dependent responses (Talavage *et al.*, 2000).

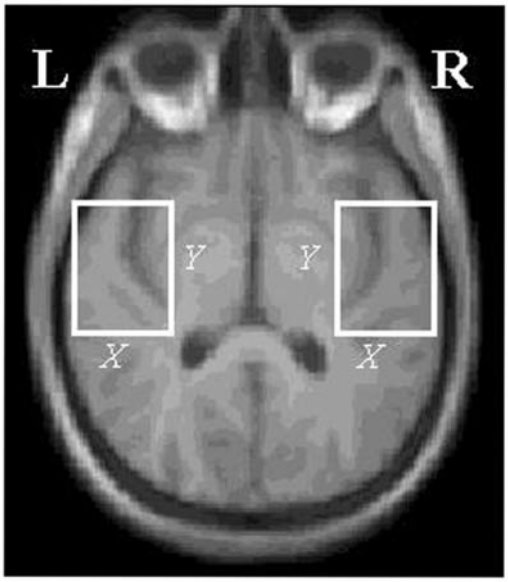

Figure 3. Selected window in the left and right auditory cortices that includes primary and non-primary areas on the STP. The image plane corresponds to that shown for individual subjects in the lower panel of Figure 4. These bilateral windows defined the area for further description of condition-specific responses across putative anatomical subdivisions of the auditory cortex.

The primary auditory field generally occupies the medial portion of the anterior-most transverse gyrus of Heschl (HG) in both hemispheres (Galaburda and Sanides, 1980) and this gyrus can be observed on structural MR scans. The size, shape and location of the primary auditory field have been described recently (Morosan *et al.*, 2001; Rademacher *et al.*, 2001). They report that the convexity of the anterior-most Heschl's gyrus appears to be largely devoted to primary auditory cortex, although it can also extend onto the planum polare and planum temporale by variable amounts across subjects. Despite this variability in extent, estimates for the centre of the primary auditory cortex appear quite stable across studies. For example, we have calculated that the centroids of the architectonically defined primary area (Rademacher *et al.*, 2001) are, on average, only 3.3 mm (one voxel) away from the centroids of the morphologically defined HG (Penhune, 1996). The centroids of either probability map therefore provide adequate estimates for the centre of primary auditory cortex.

The absolute locations of non-primary fields also vary, but their location relative to the primary field and their interrelationships appear more constant (Galaburda and Sanides, 1980). For the two human brains studied by Rivier and Clarke, five non-primary regions were identified and descriptively labelled according the their position relative to A1; the anterior area (AA), the posterior area (PA), the lateral are (LA) the medial area (MA) and the superior temporal area (STA). The central locations of these five regions are reported in standard brain space (see Table 1). This allows direct, objective cross-reference with imaging data that have been similarly transformed. For example, stable estimates for the centre of the primary area in the imaging data were calculated from the probability-weighted centre of mass of HG given by Penhune *et al.* (Penhune *et al.*, 1996) and were $x = -45$, $y = -20$, $z = 8$ mm in the left hemisphere and $x = 48$, $y = -16$, $z = 7$ mm in the right hemisphere. Using the coordinates given by Rivier and Clarke (Rivier and Clarke, 1997), the central location for each non-primary architectonic region was expressed in terms of its mean Euclidean distance in the X, Y and Z planes from this centre of the primary area. These architectonically based distance measures were then applied to the group-normalized imaging data to identify the likely central point of each region from the centre of the primary area. Using this procedure, the spatial layout of the estimated centres of each region

Downloaded from https://academic.oup.com/cercor/article/12/2/140/301066 by guest on 21 August 2022

Table 1
Estimates of the location of non-primary areas on the STP based on the distance measures of each architectonic field from the primary area applied from the centre of HG in the normalized functional data

Area	Surface area (cm^2)	Estimated radius (mm)	Left hemisphere			Right hemisphere		
			x	*y*	*z*	*x*	*y*	*z*
AA	0.95	5.5	–51.1	–1.8	5.8	56.3	–4.1	3.7
PA	0.80	5.0	–41.5	–30.3	7.8	41.8	–28.1	12.3
LA	2.98	9.7	–58.6	–25.3	6.8	62.8	–22.1	7.7
MA	2.40	8.7	–44.1	–17.3	2.3	47.8	–14.6	–1.3
STA	1.83	7.6	–63.6	–26.8	–0.2	68.3	–25.1	8.7

Mean coordinates are reported in mm (*x*, medio-lateral; *y*, antero-posterior; *z*, infero-superior) in standard brain space. Mean surface area measurements of putative auditory cortical fields were measured by Rivier and Clarke (Rivier and Clarke 1997) using flat reconstructions and the radius of each field is estimated from this value.

agreed with the schematized organization of regions relative to the gyri and sulci on the STP given by Rivier and Clarke (Rivier and Clarke 1997). As an additional step, the radius of each non-primary auditory region (assuming a spherical field) was estimated from its surface area reported by Rivier and Clarke (Rivier and Clarke 1997) for the two brains (see Table 1). The estimates of the central location and the radius of each region around that centre defined the boundaries for determining whether or not functional activation occurred within that region.

Results

Target-discrimination Performance

Subjects were highly accurate at detecting target tones; mean accuracy was 91% (SD = 13.5%) and mean *d′* was 4.80 (SD = 0.93). There was no effect of spectral or temporal complexity on the ability to detect targets [$F(1,5) = 4.59$ and 4.53, respectively, $P = 0.08$], so it is not surprising that performance-correlated cortical activation was not observed.

Imaging Data

Relative to the silent baseline, the tone stimuli generated widespread bilateral activation along the superior temporal gyrus (STG). Small clusters of activation were also seen in the precentral gyrus, which are likely to reflect the right-handed finger press in response to target stimuli. This activation was bilateral in five out of the six subjects, while the left-handed subject (who also responded using her right hand) generated only left-sided, precentral gyral activation.

Although the four tone stimuli elicited STG activation relative to the silent condition in all subjects, the extent and amplitude of this activation varied across tone conditions. The pattern of differential auditory activation more specifically identifies which brain areas show a greater response to the two main experimental effects (i.e. to the increase in frequency spectrum and to the frequency modulation). To investigate which brain areas responded to the greater bandwidth, irrespective of the modulation, a conjunction analysis identified those areas that were significantly more activated by the static harmonic complex than by the static single tone, and by the FM harmonic complex than by the FM single tone. Typically, activation foci were observed in HG in the right hemisphere and in adjacent auditory areas on the dorsolateral STG on both sides (areas shown in green and yellow in Fig. 4). According to the parcellation scheme reported by Morosan *et al.* (Morosan *et al.*, 2001), the fMRI peak in HG was located within the medial subdivision of the primary auditory cortex, Te1.1. Using a probability-weighted map derived from these 10 normalized brains (Johnsrude *et al.*, 2001), the probability that the peak occurred in Te1.1 was 0.17. Using the location of the fMRI peaks in dorsolateral STG, reported in Table 2, and the methods for estimating the central location and radius of each non-primary auditory area, we estimated that non-primary activation putatively occupied the regions LA and STA.

To locate the effects of frequency modulation, irrespective of the carrier signal, a second conjunction analysis identified brain areas that responded preferentially to FM, relative to the static tones, for both the harmonic-complex and single tone contrasts. Again, bilateral activation was observed in the STG. The peaks of activation for the frequency-modulation contrast had consistently greater percentage signal differences and *T*-scores than those for the harmonic contrast (Table 2), indicating a stronger differential response to the temporal cue. Activation encompassed many of those brain areas that showed the harmonic effect (such areas of response overlap are illustrated in green in Fig. 4), but was more widespread in all subjects (additional areas are shown in pink in Fig. 4). Evaluation of the location of fMRI peaks in non-primary auditory cortex (reported in Table 2) estimated that activation occupied regions LA and STA, in both hemispheres. Again, using the parcellation scheme of Morosan *et al.* (Morosan *et al.*, 2001), fMRI peaks along HG were located in area Te1.0 and were more lateral than the harmonic peak. The probability of this peak being in the left Te1.0 was 0.11 and in the right was 0.06. The typical locations of non-primary peak responses to FM in STA were almost identical to those showing the effect of harmonicity (being 0 and 2 mm distant in left and right hemispheres). For LA, these peaks were shifted by 4 mm (anteriorly) and 11 mm (medially) respectively from the peaks for harmonicity. Activation extended inferiorly to include the superior temporal sulcus (STS), and this was more posterior on the left side than on the right. A bilateral region of activation was also found in anterolateral STG, anterior to HG. The spatial coordinates of the peaks in this area were –68, –2, –10 on the left side and 62, 2, –6 on the right side (see Table 2). The locations of the peak in this area were 23 mm (on the left) and 13 mm (on the right) away from the estimated centres of region AA, being more anterior and lateral to AA. Given the 5.5 mm estimated radius of AA, it is unlikely that these peak responses are located within the architectonic region AA. We thus refer to this area as anterolateral STG (alSTG), to distinguish it from the architectonically defined area, AA, reported by River and Clarke (River and Clarke 1997).

The lower panel in Figure 4 shows the differential effects of the two stimulus manipulations on the pattern of STP activation for each of the six subjects. All subjects showed principal aspects of the group-average result in that harmonic effects occupied primary regions and the FM effects extended into anterolateral and dorsolateral non-primary regions. For subjects 1, 3, 4 and 6, the conjoint harmonic/FM activation (shown in yellow in Fig. 4) occurred bilaterally within portions of the HG, and, in subject 2, this occurred for the right HG. As in the previous analyses, activation was taken to involve a particular

Downloaded from https://academic.oup.com/cercor/article/12/2/140/301066 by guest on 21 August 2022

Table 2
Main effects of spectral and temporal complexity on cortical activation

Coordinates (mm)			*T*-score	% signal difference	Hemisphere	Putative auditory area
x	*y*	*z*				
Harmonic-complex versus single tone						
48	–20	6	6.71	0.34	R	HG
–54	–22	2	4.76	0.23	L	LA
66	–18	10	3.50	0.25	R	LA
–64	–14	2	4.57	0.22	L	STA
68	–6	2	4.17	0.36	R	STA
Frequency-modulated versus static tone						
–54	–14	4	8.15	0.39	L	HG
56	–10	2	11.68	0.67	R	HG
–68	–2	–10	5.44	0.38	L	alSTG
62	2	–6	6.58	0.45	R	alSTG
–54	–18	4	6.85	0.38	L	LA
56	–16	4	8.45	0.48	R	LA
–64	–14	2	10.81	0.53	L	STA
64	–6	2	10.86	0.80	R	STA
–66	–42	6	3.53	0.24	L	STS
62	–20	–2	9.95	0.41	R	STS

Coordinates and *T*-scores are reported for the peak voxels of activation located within HG and the non-primary auditory fields.

non-primary auditory area if it overlapped with the boundary that was defined by the estimated centre and radius of that area. Although the shape and absolute spatial location of the additional activation to FM (shown in pink in Fig. 4) varied across individuals, it incorporated portions of the same non-primary auditory fields in almost all six subjects. Subject 4 had widespread activation that included areas LA, STA and alSTG. While such activation was not quite so extensive for the other five subjects, portions of LA, STA and alSTG also demonstrated a greater response to the FM than to the static tones. In terms of the overall pattern of the differential activation by the tones, results for the left-handed subject (subject 4) were qualitatively similar to those for the right-handed subjects.

Condition-specific Responses in Architectonic Regions

While Figure 4 usefully summarizes the extent of significant auditory activation, information about the actual magnitude of the condition-specific response at each voxel is not represented by these thresholded figures. To address this question, a more detailed matrix of responses was generated by plotting the response magnitudes for all voxels within bilateral selected windows that passed through the STP (shown in Fig. 3).

Plots of activation for each stimulus condition were strikingly symmetrical across hemispheres and revealed spatially coherent ridges and valleys. The estimated locations of the putative subdivisions of the auditory cortex, given in Table 1, were overlaid onto these matrices for anatomical interpretation. The long axes of the ridges were in close parallel to the axis of HG, supporting a correspondence between fMRI activation and cortical anatomy. Little activation to the static single and harmonic-complex tones was observed. In contrast, strong positive responses were observed for both FM stimuli. For the FM single tone, these were predominantly in the lateral portion of HG, in a lateral region just posterior to the axis of HG that may correspond to part of the elongated region, STA and in an anterolateral region (alSTG), possibly greater on the right (Fig. 5). The FM harmonic-complex tone evoked a particularly high, localized response close to STA, with effects also observed in lateral HG and alSTG, where again the effect appeared somewhat slightly greater on the right (Fig. 5). There were no observable condition-specific effects in regions AA, MA and PA. Thus, the overall activation patterns summarized in Figure 4 conceal a detailed fine structure that is somewhat stimulus-specific.

Discussion

In the present study, effects of spectral and temporal complexity were found both in primary and non-primary regions, with more widespread activation in the non-primary regions in response to frequency modulation. As spectral complexity increased, by adding five harmonics to a single frequency tone, activation increased in primary auditory fields and in portions of adjacent non-primary fields, LA and STA, in dorsolateral STG. Similarly, as temporal complexity increased, by temporally varying the frequency components in the tone, activation also increased in the above auditory regions. Activation induced by these temporal cues was stronger than that induced by spectral cues, as indicated by higher magnitude signal differences and *T*-scores. We found no clear spatial segregation of responses to either spectral or temporal cues in these three regions, only differences in the magnitude of activity. The additive effects of spectral and temporal complexity in parts of HG, LA and STA may indicate a general increase in activation with increasing acoustical complexity in these regions. Lack of spatial segregation might arise from overlapping or complex response fields to both spectral and temporal patterns, although the response fields cannot be simply defined by the specific acoustical manipulations in the present study. Response selectivity to the FM tones was observed in additional auditory regions, namely the STS and in an area on the STP that was anterior and lateral to AA (alSTG), while the most complex stimulus (FM harmonic-complex tone) appeared to elicit a high peak of activation lateral and just posterior to the axis of HG, possibly in a region of STA.

Thus, in summary, we suggest that the spectral and temporal cues used in the present study activate auditory regions in common, but that temporal cues are more salient than spectral cues and that their processing can involve additional non-primary auditory regions, such as anterolateral cortex.

Downloaded from https://academic.oup.com/cercor/article/12/2/140/301066 by guest on 21 August 2022

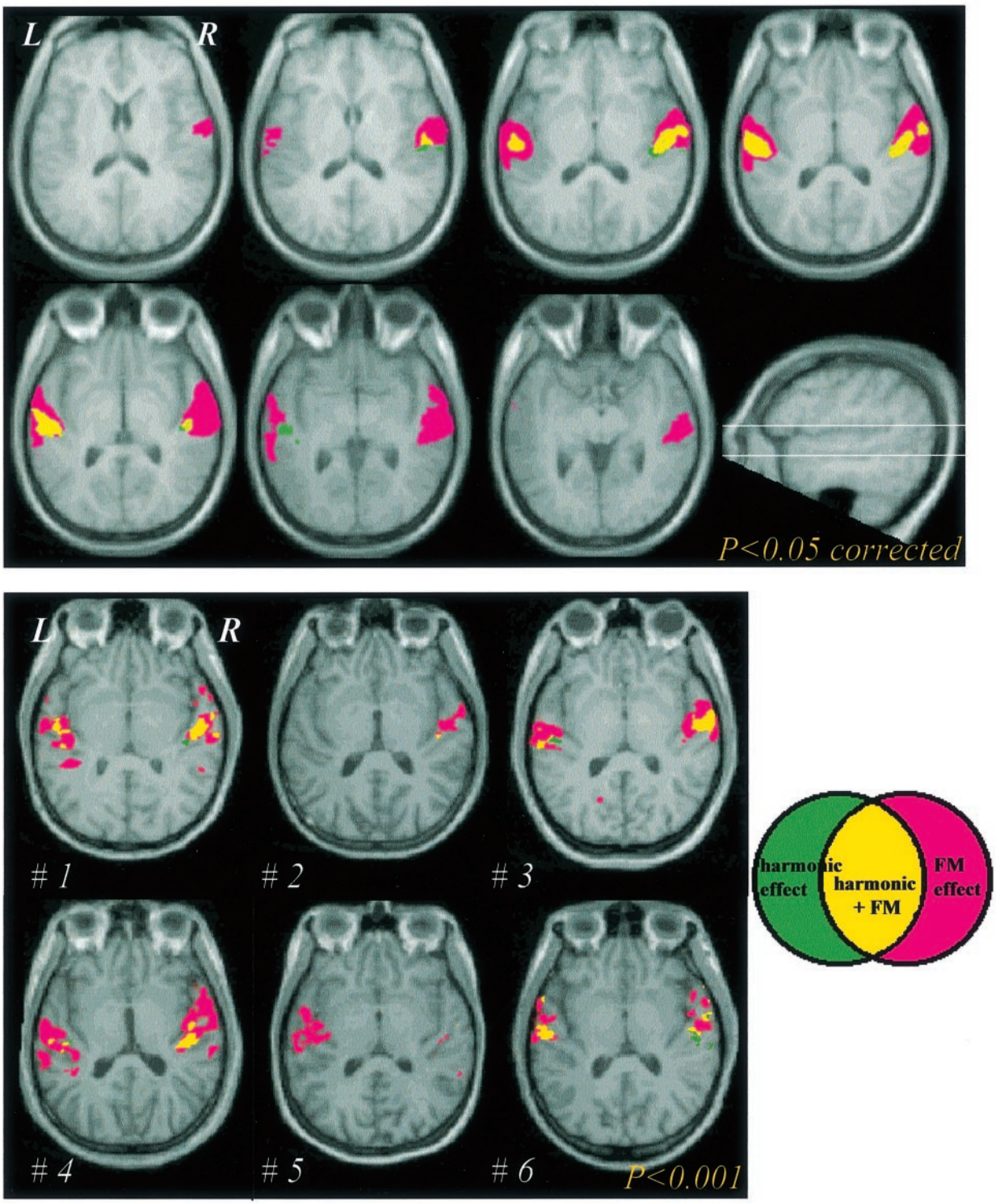

Figure 4. Oblique–axial slices parallel to the STP showing regions of activation produced by auditory stimulation. Slices are displayed in neurological convention. The upper panel displays the *T* map for the group-wise analysis. Activation is overlaid onto the mean structural image viewed at 4 mm intervals, from the STP down to the STS. The sagittal view illustrates the slice orientation selected for display. The lower panel presents *T* maps for the six individual subjects that are overlaid onto individual structural scans. These *post hoc T* maps are thresholded at $P < 0.001$, uncorrected for multiple tests across the whole brain volume. The image plane corresponds to the STP and corresponds approximately midway between the slices 3 and 4 that are shown in the above panel. The key at the bottom of the figure indicates the nature of the auditory activation. Green indicates brain areas more activated by harmonic tones than by single tones. Pink indicates those regions that were more activated by frequency-modulated tones than by static tones. Yellow indicates those regions showing an additive (i.e. combined) effects of both harmonicity and frequency modulation.

Relationship Between Response Peaks and Architectonic Regions in Human Auditory Cortex

According to the probabilistic map of the primary area (Penhune *et al.*, 1996), peak responses to temporal complexity were located in the lateral portion of HG, 11 mm away from its centre of mass. Additionally, from the work of Morosan *et al.*

Downloaded from https://academic.oup.com/cercor/article/12/2/140/301066 by guest on 21 August 2022

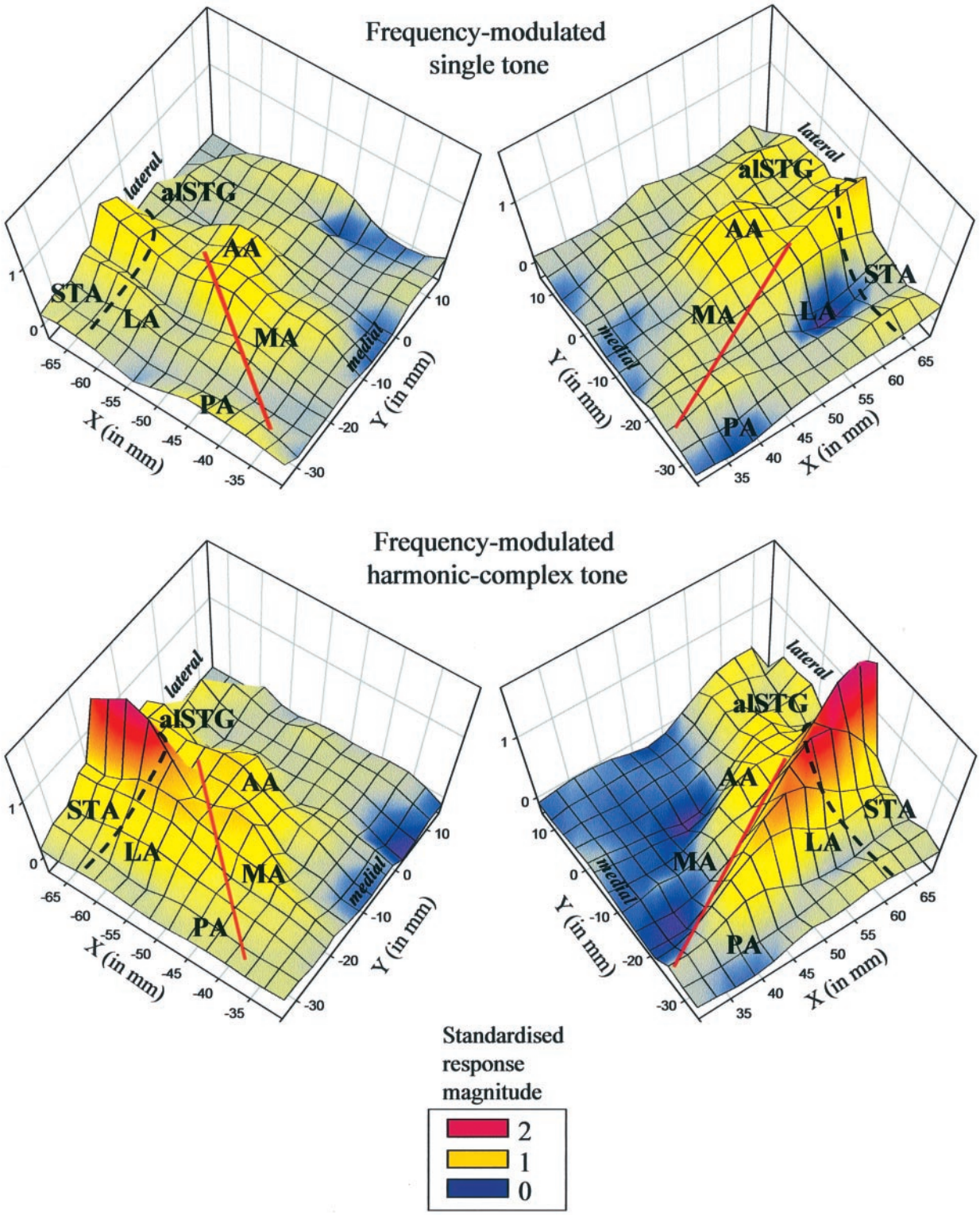

Figure 5. Response magnitudes for the frequency-modulated tones in the left and right auditory cortices. The views are displayed in neurological convention (i.e. left hemisphere is shown in the left panel) with lateral and medial zones labelled. The orientation of the long axis of HG is plotted as a red line, where the end points are defined by the 50–75% probability contour for HG (Penhune *et al.*, 1996). Approximate central locations of the surrounding architectonic fields are also depicted and the approximate borders of region STA are depicted by the dashed lines. Upper figures show the response preferences for the frequency-modulated single tone in anterolateral and posterolateral auditory regions. Lower figures for the frequency-modulated harmonic-complex tone reveal a particularly high bilateral posterolateral response that may overlap with part of STA, an elongated anatomical region along the lateral convexity of the STG.

(Morosan *et al.*, 2001), this may correspond to the architectonic subdivision of the primary area, Te1.0. In our group-averaged structural image, the peak response in HG in both hemispheres was located on its anterolateral portion. This position is at least two-thirds of the length of HG from the medial-most end of the gyrus and coincides with the lateral

Downloaded from https://academic.oup.com/cercor/article/12/2/140/301066 by guest on 21 August 2022

extent of the primary, koniocortical field, Kalt (Galaburda and Sanides, 1980).

We have proposed a correspondence between the condition-specific response peaks and anatomically defined auditory regions based on the proximity of the peak to the estimated locations and extents of architectonic regions on the STP as given by Rivier and Clarke (Rivier and Clarke, 1997). From these estimates of the location of architectonic regions, increased activation by the spectral or temporal cues used here did not occur in three of the architectonically defined non-primary auditory areas, AA, MA and PA. In contrast, three regions of condition-specific responses were observed on the lateral part of the STP. The anterior-most region lay on the planum polare, in an anterolateral position relative to HG. We suggest that its centre was too distant from the centre of region AA for the two areas to be coincident and have therefore referred to this region as anterolateral STG (alSTG). Relating this activated region to other cytoarchitectonic reports, the position of alSTG may be close to the anterior borders of the belt of non-primary parakoniocortex [fields PaAe and PaAi (Galaburda and Sanides, 1980)], which extend quite anteriorly towards the temporal pole. PaAe appears more externally (towards the lateral convexity) than PaAi and shares a major portion of the STG and so, at its most anterior part, may overlap with alSTG. The second lateral region was immediately posterolateral to the axis of HG and, from the mean structural scans, was placed near the sulcus just behind HG. The peak of this area fell within the estimated boundaries of LA. The third region, STA, was on average 13 mm more anterolateral than LA and was located on the convexity of the STG. In the terminology of Galaburda and Sanides (Galaburda and Sanides, 1980), parakoniocortical borders extend posteriorly beyond HG to the border with the temporoparietal area, Tpt. Thus, there are possible correspondences between LA and the posteromedial portion of the internal parakoniocortex, PaAi and between STA and the external parakoniocortex, PaAe.

Prior evidence suggests that an orderly spatial segregation between response selectivity in these areas for spectral and temporal cues should not necessarily be expected. Rivier and Clarke (Rivier and Clarke, 1997), for example, conducted a meta-analysis of 10 neuroimaging studies in which they compared the reported location of the greatest peak of activation with their architectonic classification of the primary auditory region and areas AA, PA, LA, MA and STA. Although spectro-temporal features of the stimuli varied widely (e.g. noise, environmental sounds, words, music) and the stimulus contrasts were not necessarily tightly controlled in the spectro-temporal dimension, this study is important because it indicates that, whilst peaks occurred within many of these auditory areas, there was no clear systematic segregation of the responses to any of the different types of acoustical stimuli across auditory regions. It also highlights the view that the functional roles of the auditory cortical fields are poorly understood. This lack of spatial segregation is clearly compatible with our finding that, in LA and STA, the fMRI peaks for the two manipulations of spectral and temporal cues were not systematically displaced from one another. Since alSTG was not mapped by Rivier and Clarke, speculation about its response properties is not relevant here.

Response Selectivity to Spectral Complexity

The spectral identity of a sound may be encoded in the elements of the tonotopic array that are activated by it; with harmonic-complex tones activating many more frequency channels than do single tones. Neuromagnetic (Pantev *et al.*, 1989) and microelectrode (Howard *et al.*, 1996) data indicate that acoustic frequencies separated by one to two octaves generate foci of activity separated by ~6–8 mm. Therefore, the 2.6 octave-wide frequency spread of the harmonics used here would generate considerably wider activation than the single tone, within fields that are tonotopically organized. Functional MRI data suggest that at least four tonotopically organized areas exist in human auditory cortex, including portions of the primary area, LA and STA (Talavage *et al.*, 2000). Thus, the effect of the harmonic-complex tone observed in HG, LA and STA may in part be explained by the response characteristics of these functional fields that are tonotopically organized.

In primary auditory cortex, horizontal connections between neurons tuned to harmonically related frequencies may also facilitate responsiveness to the harmonic-complex tone (Kadia *et al.*, 2000). The harmonic complex and single tone also differ in terms of their perceived pitch strength, with the harmonic complex generating a stronger sense of pitch than the single tone. Increasing pitch strength has been reported to produce small foci of activation in the region of HG (Griffiths *et al.*, 1998). However, this localized sensitivity to pitch strength is insufficient to account for the spectral effects reported here.

Response Selectivity to Temporal Complexity

Analyses revealed a bilateral network of cortical sites involved in the processing of temporal complexity, including the lateral part of the primary area, the lateral portion of the STP (LA and STA), alSTG and STS. Bilateral responses to complexity in the lateral portion of the primary area suggest functional segregation within the primary area in humans. This has been further quantified in more detail (Johnsrude *et al.*, 2001), but the exact nature of this segregation is unclear. Studies of the laminar organization of the human primary field reveal either two (KAm and KAlt) or three [Te1.0, 1.1 and 1.2 (Morosan *et al.*, 2001)] anatomical sub-divisions. Nevertheless, the mediolateral distinction observed here in the functional data is consistent with there being at least two tonotopic primary fields along the long axis of HG (Talavage *et al.*, 2000).

The particularly strong response to the FM harmonic-complex tone in lateral STP indicates sensitivity to temporal and spectral structure. The responsiveness to spectro-temporally dynamic sounds in auditory fields beyond primary cortex is well-established in primates (Rauschecker, 1998a,b). For example, where the analysis of cues relevant to auditory pattern identification, such as for monkey calls, arises in anterolateral belt areas. The anterolateral pathway for speech comprehension somewhat falls within this theoretical framework since, to humans, speech is a con-specific vocalization in the same way that monkey calls are to primates. Although clearly differing from speech, the FM tones used in the present study nevertheless share the same fundamental property as speech in terms of their variation over time in frequency. Thus, the recruitment of the anterolateral cortex (alSTG) in the analysis of FM is at least consistent with the functional role of the anterolateral pathway that has been proposed by Rauschecker. In terms of the dorsolateral activation by FM, several lines of corroborating evidence in humans also indicate that this region is highly activated by specific spectro-temporal modulations. For example, electrode recordings from the surface of the superior temporal gyrus in awake humans undergoing surgery for intractable epilepsy have revealed a posterolateral region whose evoked responses discriminate between specific classes of consonant–vowel stimuli (Jenison *et al.*, 2001). Thus, this posterolateral field may have a role in speech waveform discrimination. Another imaging study has also implicated the dorsolateral surface of STG in temporal

Downloaded from https://academic.oup.com/cercor/article/12/2/140/301066 by guest on 21 August 2022

processing as this area responds more strongly to a continuous sequence of single tones that change in frequency over time (a type of FM) than to white noise (Binder *et al.*, 2000). Relative to the former single tone stimulus, words, pseudowords and reversed words also generate more activation in bilateral portions of STG lateral to the anterior aspect of HG. However, these foci tended to be more ventrally located, proximal to STS (Binder *et al.*, 2000). In the present study, activated parts of STS showed a discrete preference for temporal complexity, as did alSTG. Such anterior activation with modulation concurs with previous imaging data (Hall *et al.*, 2000b). In macaque and rhesus monkey, more distant cortical regions in the STG, the STS and prefrontal cortex are connected to belt fields (Pandya, 1995; Kaas and Hackett, 1999; Romanski *et al.*, 1999a,b) and both the anterior temporal lobe and dorsal bank of the STS receive afferent input from adjacent parabelt auditory cortex on the dorsal surface of the STG (Pandya, 1995; Hackett *et al.*, 1998), suggesting a role for these areas in higher-level acoustical or multi-modal analysis. Furthermore, neuroimaging studies (Binder *et al.*, 2000; Scott *et al.*, 2000) suggest a dorsal–ventral–anterior pathway for auditory pattern analysis; with processing of phonetic and dynamic pitch variation occurring on the dorsal STG, speech sensitivity in STS and (when stimuli are matched for acoustical complexity) speech intelligibility in left anterior STS. Our data highlight a special role for lateral primary and lateral auditory areas on the STP in the analysis of temporal complexity that may form the early cortical stages in this hierarchy.

Notes

MR scanning was provided by the Wellcome Department of Cognitive Neurology, London. Dr Ingrid Johnsrude was supported by a Wellcome Fellowship at the Wellcome Department of Cognitive Neurology, London. The other authors were supported by the Medical Research Council. We wish to thank Oliver Josephs for his helpful comments and advice on MR image acquisition, Jesper Anderson for his assistance in applying the mutual information algorithm and Miguel Gonçalves for some of the SPM99 data analysis. We also wish to thank Mark Wessinger (Wessinger *et al.*, 2001, Fig. 2), whose method of data presentation we adopt in Figure 4.

Address correspondence to Deborah A. Hall, MRC Institute of Hearing Research, University Park, Nottingham NG7 2RD, UK. Email: d.hall@ih.mrc.ac.uk.

References

Binder JR, Frost JA, Hammeke TA, Bellgowan PSF, Springer JA, Kaufman JN, Possing ET (2000) Human temporal lobe activation by speech and nonspeech sounds. Cereb Cortex 10:512–528.

Burkhard MD, Sachs RM (1975) Anthropometric manikin for acoustic research. J Acoust Soc Am 58:214–222.

Friston KJ, Holmes AP, Worsley KJ (1999) How many subjects constitute a study? NeuroImage 10:1–5.

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

Griffiths TD, Büchel C, Frackowiak RSJ, Patterson RD (1998) Analysis of temporal structure in sound by the human brain. Nat Neurosci 1:422–427.

Hackett TA, Stepniewska I, Kaas JH (1998) Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. J Comp Neurol 394:475–495.

Hall DA, Summerfield AQ, Goncalves MS, Foster JR, Palmer AR, Bowtell RW (2000a) Time-course of the auditory BOLD response to scanner noise. Magn Reson Med 43:606–606.

Hall DA, Haggard MP, Akeroyd MA, Summerfield AQ, Palmer AR, Elliott MR, Bowtell RW (2000b) Stimulus modulation and task effects in auditory processing measured using fMRI. Hum Brain Mapp 10:107–119.

Heil P, Irvine DRF (1998) Functional specialisation in auditory cortex: responses to frequency-modulated stimuli in the cat's posterior auditory field. J Neurophysiol 79:3041–3059.

Howard MA, Volkov IO, Abbas PJ, Damasio H, Ollendieck MC, Granner MA (1996) A chronic microelectrode investigation of the tonotopic organisation of human auditory cortex. Brain Res 724:260–264.

Jenison RL, Reale RA, Brugge JF, Hind JE, Bakken H, Volkov IO, Howard MA (2001) Information-theoretic maps of speech evoked potentials recorded directly from human auditory cortex, #663. Proceedings of the 24th Midwinter Meeting of the Association for Research in Otolaryngology, p. 188.

Johnsrude I, Cusack R, Morosan P, Hall D, Brett M, Zilles K, Frackowiak R (2001) Cytoarchitectonic region-of-interest analysis of auditory imaging data. NeuroImage 13:S897.

Kaas JH, Hackett TA (1999) 'What' and 'where' processing in auditory cortex. Nat Neurosci 2:1045–1047.

Kadia S, Snider R, Wang X (2000) Influence of stimulus components placed outside classical receptive field reveals harmonic structure of the auditory cortex. Proceedings of the 23rd Midwinter Meeting of the Association for Research in Otolaryngology, pp. 14–15.

Liegeois-Chauvel C, Musolino A, Chauvel P (1991) Localization of the primary auditory area in man. Brain 114:139–153.

Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P (1997) Multimodality image registration by maximization of mutual information. IEEE Trans Med Imaging 16:187–198.

Merzenich MM, Brugge JF (1973) Representation of the cochlear partition on the superior temporal plane of the macaque monkey. Brain Res 50:275–296.

Moore BCJ, Glasberg BR, Baer T (1997) A model for the prediction of thresholds, loudness and partial loudness. J Audio Eng Soc 45: 224–240.

Morel A, Garraghty PE, Kaas JH (1993) Tonotopic organization, architectonic fields, and connections of auditory-cortex in Macaque monkeys. J Comp Neurol 335:437–459.

Morosan P, Rademacher J, Schleicher A, Amunts K, Schormann T, Zilles K (2001) Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. NeuroImage 13:684–701.

Moutoussis K, Zeki S (1997) Functional segregation and temporal hierarchy of the visual perceptive systems. Proc R Soc Lond B Biol Sci 264:1407–1414.

Palmer AR, Bullock DC, Chambers JD (1998) A high-output, high-quality sound system for use in auditory fMRI. NeuroImage 7:S359.

Pandya, DN (1995) Anatomy of the auditory cortex. Rev Neurol (Paris) 151:486–494.

Pantev C, Hoke M, Lütkenhöner B, Lehnertz K (1989) Tonotopic organisation of the auditory cortex: pitch versus frequency representation. Science 246:486–488.

Penhune VB, Zatorre RJ, Macdonald JD, Evans AC (1996) Interhemispheric anatomical differences in human primary auditory-cortex: probabilistic mapping and volume measurement from magnetic resonance scans. Cereb Cortex 6:661–672.

Price CJ, Friston KJ (1997) Cognitive conjunction: a new approach to brain activation experiments. NeuroImage 5:261–270.

Rademacher J, Caviness VS, Steinmetz H, Galaburda AM (1993) Topographical variation of the human primary cortices — implications for neuroimaging, brain mapping, and neurobiology. Cereb Cortex 3:313–329.

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 JP (1998a) Parallel processing in the auditory cortex of primates. Audiol Neurootol 3:86–103.

Rauschecker JP (1998b) Cortical processing of complex sounds. Curr Opin Neurobiol 8:516–521.

Rauschecker JP, Tian B (2000) Mechanisms and streams for processing of 'what' and 'where' in auditory cortex. Proc Natl Acad Sci USA 97:11800–11806.

Rauschecker JP, Tian B, Hauser M (1995) Processing of complex sounds in the macaque non-primary auditory cortex. Science 268:111–114.

Rivier F, Clarke S (1997) Cytochrome oxidase, acetylcholinesterase, and NADPH-diaphorase staining in human supratemporal and insular cortex: evidence for multiple auditory areas. NeuroImage 6:288–304.

Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP (1999a) Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nat Neurosci 2:1131–1136.

Downloaded from https://academic.oup.com/cercor/article/12/2/140/301066 by guest on 21 August 2022

Romanski LM, Bates JF, Goldman-Rakic PS (1999b) Auditory belt and parabelt projections to the prefrontal cortex in rhesus monkey. J Comp Neurol 403:141–157.

Scheich H (1991) Auditory cortex: comparative aspects of maps and plasticity. Curr Opin Neurobiol 1:236–247.

Schreiner CE, Cynader MS (1984) Basic functional organisation of secondary auditory cortical field (AII) of the cat. J Neurophysiol 51:1284–1305.

Schreiner CE, Urbas JV (1988) Representation of amplitude-modulation in the auditory-cortex of the cat. 2. Comparison between cortical fields. Hear Res 32:49–64.

Scott SK, Blank SC, Rosen S, Wise RJS (2000) Identification of a pathway for intelligible speech in the left temporal lobe. Brain 123:2400–2406.

Talairach J, Tournoux P (1988) Co-planar stereotaxic atlas of the human brain. Stuttgart: Thieme.

Talavage TM, Ledden PJ, Benson RR, Rosen BR, Melcher JR (2000) Frequency-dependent responses exhibited by multiple regions in human auditory cortex. Hear Res 150:225–244.

Tian B, Rauschecker JP (1994) Processing of frequency-modulated sounds in the cat's anterior auditory field. J Neurophysiol 71: 1959–1975.

Tian B, Rauschecker JP, (1998) Processing of frequency-modulated sounds in the cat's posterior auditory field. J Neurophysiol 79:2629–2642.

Wallace MN, Rutkowski RG, Palmer AR (1999) A ventrorostral belt is adjacent to the guinea pig primary auditory cortex. NeuroReport 10:2095–2099.

Wessinger M, VanMeter J, Tian B, Van Lare J, Pekar J, Rauschecker JP (2001) Hierarchical organisation of the human auditory cortex revealed by functional magnetic resonance imaging. J Cogn Neurosci 13:1–7.

Downloaded from https://academic.oup.com/cercor/article/12/2/140/301066 by guest on 21 August 2022