Functional Magnetic Resonance Imaging Assessment of the Human Brain Auditory Cortex Response to Increasing Word Presentation Rates

AJAY DHANKHAR, 4 BRUCE E. WEXLER, ROBERT K. FULBRIGHT, TERRY HALWES, ANDREW M. BLAMIRE, 3 AND ROBERT G. SHULMAN 4

1Departments of Psychiatry, 2Radiology, 3Neurosurgery, and 4Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, Connecticut 06510

Dhankhar, Ajay, Bruce E. Wexler, Robert K. Fulbright, Terry Halwes, Andrew M. Blamire, and Robert G. Shulman. Functional magnetic resonance imaging assessment of the human brain auditory cortex response to increasing word presentation rates. J. Neurophysiol. 77: 476–483, 1997. In an investigation of the auditory cortex response to speech, six subjects were studied using echo-planar functional magnetic resonance imaging (fMRI) at 2.1T. The subjects were asked to listen to English nouns presented at various rates ranging from 0 words per minute (wpm) to 130 wpm while fMRI images encompassing their primary and posterior superior secondary auditory cortices were acquired. An asymmetric spin echo imaging sequence was used with an induced T2 weighting of 50 ms to allow for transverse relaxation effects. Images were acquired in two or four axial-oblique slices with a repetition time of 3.75 or 7.5 s, in plane resolution of 6 × 3 mm, and a slice thickness of 5 mm. Localized activation centered over grey matter was consistently observed in all subjects in the transverse temporal gyrus (TTG), the transverse temporal sulcus (TTS), and the posterior superior aspect of the superior temporal gyrus (STG). The total activated volume and the integrated signal response in bilateral primary and posterior superior secondary auditory cortices increased with increasing rate of word presentation, peaking at 90 wpm (with some intersubject variability) with a subsequent fall at 130 wpm. There were no significant differences in the rate dependence of the signal response in bilateral primary and bilateral posterior superior secondary auditory cortices (P < 0.05).

INTRODUCTION

Auditory processing in the human brain has been investigated extensively using several different stimuli including tones, syllables, and words. The array of techniques include cytoarchitectural mapping (Galaburda and Sanides 1980), clino-pathological correlates of patients with lesions or auditory deficiencies (Benson 1979), open brain electrical stimulation recordings (Penfield and Rasmussen 1950), evoked potentials (Celesia 1976; Hari 1991) positron emission tomography (PET) (Lauter et al. 1985; Mazziotta et al. 1982; Petersen et al. 1988; Wise et al. 1991) and functional magnetic resonance imaging (fMRI) (Binder et al. 1994a,b). Although anatomic boundaries between different auditory processing areas are not precisely known, and may vary among individuals, it is generally well recognized that the primary auditory cortex is located on the transverse temporal gyrus (TTG or Heschl’s gyrus), which is Brodmann’s Area (BA) 41, while the secondary auditory (or association) cortex is in surrounding regions on the superior temporal gyrus (STG) encompassing Brodmann’s Areas 21, 22, 42, and 52 (Celesia 1976; Talairach and Tournoux 1988).

This understanding has been strengthened recently by PET studies using baseline noise, tones, words, and speech that have shown similar bilateral activations within the primary cortices (Lauter et al. 1985; Mazziotta et al. 1982; Petersen et al. 1988; Wise et al. 1991). However, speech stimuli produced larger and more widespread activations posteriorly and superiorly and activated the left hemisphere to a greater degree than the right. This was speculated to be related to the specialized role of the left hemisphere in semantic processing (Wise et al. 1991; Zatorre et al. 1992). Recently, Binder et al. (1994a) have demonstrated the feasibility of using MRI to perform functional studies of auditory processing despite a constant background scanner noise. Using words, non-words, text, and white noise as stimuli, they report localized activation patterns in the TTG and the STG consistent with previous PET studies.

In further investigation of auditory cortex activation, Price et al. (1992) used PET to compare blood flow changes associated with different rates of stimulus presentation. They instructed subjects to listen passively while single-syllable nouns were presented at rates of 10, 30, 50, 70, and 90 words per minute (wpm). When they searched for voxels that showed linear increases in blood flow across the range of stimulation rates, they found that such voxels were present in bilateral primary auditory areas and the secondary auditory area in the right hemisphere. However, a search they conducted for voxels where the regional cerebral blood flow (rCBF) difference between rest and 10 wpm was significantly greater than the average rCBF difference between each word presentation rate yielded voxels located in, and limited to, the left posterior secondary auditory cortex.

They suggested that the absence of rate dependent activation in the left posterior secondary auditory cortex was associated with its special role in processing language stimuli and that the examination of the relationship between stimulation rate and activation may be a method for identifying brain regions with distinct functional characteristics. A recent fMRI study, however, found rate-dependent increases in signal in the left posterior secondary auditory cortex that were essentially the same as those in the right posterior secondary auditory cortex and in bilateral primary auditory cortices (Binder et al. 1994b).
Methodological differences between the two studies limit comparison. The previous fMRI study used nonsense syllables as stimuli rather than words and required subjects to actively discriminate among the stimuli rather than listen passively. Moreover, the fMRI study employed a different data analysis strategy, identifying the eight voxels in each temporal lobe showing the greatest responses at the fastest stimulation rate and then looking at rate dependence of response in these voxels. This approach introduces a selection bias for voxels showing higher activation at the highest stimulation rate and, because most of the eight voxels were in primary cortices, provides only a limited sampling of the left posterior auditory areas.

Given the significance of establishing physiological differences between primary sensory and specialized association cortex, we have used speech stimuli (monosyllable English nouns) and a stimulation protocol closely paralleling that used in the PET study of Price et al. (1992). Using fMRI, we wished to characterize the changes in regional brain activation with increased processing load. In addition, we were interested particularly in clarifying the differences, if any, in the rate dependence of activation in the primary and the posterior secondary auditory cortices. A preliminary report on this research has been presented earlier (Dhandkar et al. 1994).

METHODS

Experimental setup

Six subjects (4 female, 2 male) were studied at 2.1T on a modified Bruker Biospec I (Bruker, Billerica, MA) whole-body spectrometer equipped with active shielded gradient coils (Oxford Magnet Technology, Oxford, UK). One female subject was studied twice to examine the reproducibility of the results. All subjects were right-handed, defined as writing and doing at least five of seven everyday tasks with their right hands (Wexler and Halwes 1985). All were native English speakers. None had any prior history of hearing deficits or neurologic or psychiatric illness. All subjects gave informed written consent for the experiment as approved by the Human Investigation Committee of the Yale University School of Medicine. Subjects performed the experiment voluntarily and were paid an hourly stipend as compensation.

Subjects were positioned supine in the magnet. To reduce movement during the experiment, the subject’s head was held in place by a cushioned, foam rubber holder encased in a plastic helmet fastened tight by velcro strips. The head holder was inserted inside a linear birdcage RF coil with the back of the head positioned 5 cm below the isocenter of the magnet.

For stimulus presentation, subjects wore padded, plastic shell headphones with sound blocking ear muffs to reduce background noise from the scanner. Prerecorded stimuli were played on a Nakamichi cassette deck through a high quality speaker enclosed in a specially constructed plastic box with a conical opening of 1.25-in diameter. Flexible tubing was attached to this opening and run into the magnet room where it forked into extension tubes to each ear. The intensity of the stimulus coming through the padded headphones was adjusted to a level that subjects felt was comfortably above the background scanner noise. Amplifier settings necessary to achieve this level were highly consistent across subjects. When measured in two subjects, the level was found to be 106 and 108 dB SPL. These same subjects reported that words became indistinguishable from the bursts of machine noise when stimulus intensity was reduced to 93 dB SPL for one subject and 80 dB SPL for the other subject. Thus experimental stimulus levels were substantially above the discrimination threshold subjectively and quantitatively.

Imaging protocol

Four contiguous, multislice, T1-weighted sagittal images (image matrix = 128 × 128; slice thickness = 5 mm; inversion time TIR = 750 ms; echo time TE = 18 ms; repetition time TR = 2.2 s) were acquired through the midline of the brain (the interhemispheric sagittal plane). The coordinates of the anterior commissure (AC) and the posterior commissure (PC) used in the standard Talairach-Tournoux coordinate system were noted using these images (Talairach and Tournoux 1988). T1-weighted axial-oblique images parallel to the AC-PC line and perpendicular to the midline were then used to localize two (in 4 subjects) or four (in 2 subjects) contiguous slices that included the transverse temporal gyrus (TTG or Heschl’s gyrus) and the posterior-superior aspect of the STG. In the two subjects for whom four slices were obtained, we also imaged the STG inferior and anterior to the TTG. The slices typically spanned a region 0–30 mm above the AC-PC line.

T2-weighted images (TE = 50 ms) were acquired to gauge the magnetic field homogeneity in the region of interest (ROI). To maximize task related signal changes, magnetic field homogeneity in the ROI was improved in some cases using an autoshim routine that adjusted all first and second order shim coils (Gueret and Boesch 1992). This was in addition to nonlocalized hand shimming of first order coils. Two-dimensional time of flight MR angiograms were acquired (TE = 40 ms) in all cases to localize large blood vessels in each selected slice.

Functional images were acquired using an asymmetric spin echo version of the echo planar imaging (EPI) sequence. A T2 weighting of 50 ms was induced by offsetting the spin echo from the centre of acquisition (Blamire et al. 1992). The multislice EPI images were acquired with a TR of 3.75 s (2 acquired slices) or 7.5 s (4 acquired slices), an image matrix of 64 × 64 with a nominal inplane resolution of 6 × 3 mm, and a slice thickness of 5 mm. Four dummy scans per slice were acquired prior to data collection to achieve steady-state magnetization.

Tasks

Stimuli were single syllable, common English nouns recorded by a male speaker. Words that were verbs as well as nouns (e.g., box, map) were excluded unless their use as a verb was far less common than their use as a noun. No word was presented twice. Stimuli were presented at 0 (resting), 10, 50, 90, and 130 wpm in 1-min stimulation blocks. Two or three blocks were presented at each rate, with the first presentation in ascending order of frequency (0, 10, 50, 90, 130 wpm), the
second presentation in descending order, and the third presentation in random order (50, 90, 0, 130, 10 wpm).

Although all words were monosyllabic, the average duration of utterances decreased as the presentation rate increased (715, 680, 576, and 424 ms, respectively). However, the total duration of auditory stimulation during each 1-min block increased steadily (7, 34, 52, and 55 s, respectively), albeit nonlinearly. At the fastest rate of 130 wpm, the stimuli were essentially continuous. The fundamental frequency (172 Hz) and the spectral range (172 Hz to 9 kHz) were the same at all rates. The distribution of power across frequencies changed from word to word but not from rate to rate.

Sixteen images were acquired for each slice. Subjects were instructed to rest during images 1–4 (pretask or baseline), listen attentively to the presented stimuli during images 5–12 (task), and rest again during images 13–16 (posttask).

Data analysis

LOCALIZATION OF ACTIVATION. Data analysis was performed off-line on a Vaxstation 3200 computer (Digital Equipment, Maynard, MA) equipped with a SkyWarrior array processor (Sky Computers, Lowell, MA). For each set of 16 images in a single subject, the static baseline was removed by subtracting the mean of the four pretask images ($S_p$) on a voxel by voxel basis from every image in the series. This generated a new series of 16 difference images ($\Delta S$) revealing task-related intensity changes. Statistical maps based on paired Student’s t-test comparisons between the pretask and the task periods were created as previously described (McCarthy et al. 1993). The first image acquired during stimulus presentation in every slice was not included in the t-test calculation, allowing for a 3.75-s delay in the rise of the fMRI signal (Blamire et al. 1992). T-images of the repeated sets for each frequency were averaged together, resulting in the final t-image for a performed task for each subject. The final t-images were thresholded at a t value of 1.83 ($P < 0.05, 9 df$). The method used here has the relative advantage over other similar methods (Blamire et al. 1992; Shaywitz et al. 1995) in that it involves no filtering, cluster-analysis or other user-modulated image processing routines. Minimal postprocessing manipulation of data was required for the subsequent comparisons that had to be made in this study. Anatomic localization was performed by linearly interpolating the final t-test image to a matrix size of $128 \times 128$ and superimposing it on the T1-weighted scout image of the same slice. Activated pixels are depicted on a color scale.

To quantify differences in activation between the primary and the posterior secondary auditory cortices, we segregated these two areas in each hemisphere by correlating sulcal and gyral anatomy seen on the sagittal and axial-oblique T1-weighted images. The primary auditory cortex was defined on the TTG (BA 41 or Heschl’s gyrus) by identifying one or two short gyri of the superior surface of the STG that originate from the posterior insular area and that follow an obliquely anterior course to the lateral aspect of the STG. Included in the primary auditory cortex and forming its posterior margin was the transverse temporal sulcus (TTS), which runs immediately posterior to and parallel with the TTG. Activations in the TTS therefore were considered within the primary auditory cortex. We defined posterior secondary auditory cortex (BA 22 and 42) as that aspect of the posterior and superior STG delineated anteriorly by the TTS and extending posteriorly to include the ascending posterior segment of the STS (Bogen and Bogen 1976; Cælesio 1976; Galaburda and Sanides 1980; Steinmetz and Seitz 1991; Talairach and Tournoux 1988). This was done individually for every subject, thereby taking into account intersubject variability in anatomy. For each of the four segregated anatomic regions (2 in each hemisphere), we counted significantly activated voxels ($P < 0.05$) at each rate; calculated time courses of activation, the percent $\Delta S/S_o$, and the average percent $\Delta S/S_o$ (images 6–12) during activation of the significantly activated voxels; calculated the integrated signal change within a region, which we defined as the product of the average percent $\Delta S/S_o$ and the number of significantly activated voxels; and calculated the Talairach-Tournoux coordinates of the center of the region.

Because our anatomic and functional MR images were acquired in a plane parallel to the AC-PC line and perpendicular to the midline of the brain, they were well suited for calculation of Talairach-Tournoux coordinates. This was done for each subject by stretching and coregistering the T1-weighted axial-oblique slices to match a standard anatomic image selected from this study. The selected image most closely matched the standard Talairach brain in size and anatomic configurations. From this coregistration, the Talairach coordinates of the four segregated regions were determined using home-written software.

MOTION. All acquired EPI data were subjected to a three-step examination for motion contamination. 1) Image series were viewed as a movie to detect movement of the perimeter of the head. 2) Center-of-mass (COM) algorithm: the image series were recreated by setting the intensities of all the voxels outside the brain to 0. Hence, artifactual fluctuations of voxels outside the brain did not influence our investigation for COM deviations. The COM of every image in the recreated series was calculated. Using the first image in the series as a standard, the $(x, y)$ deviation of the COM for the subsequent images was plotted against time to reveal the temporal pattern of motion. Image series showing a COM shift 30% of a voxel in $x$ or $y$ directions were discarded. Particular attention was paid to COM shifts coinciding with the onset and/or the cessation of the task, where an even more stringent standard of $>20\%$ shift in $x$ or $y$ was enforced. 3) Activated regions that showed persistence of response after stimulus cessation were treated with suspicion. Data sets that consistently produced activations showing such behavior were discarded.

As a result, four stimulation blocks were discarded from one subject, two stimulation blocks from another subject, and two other subjects had a single data block discarded. At least two blocks were available for each rate for every subject.

RESULTS

Brain mapping

Localized activations in the primary and secondary auditory cortices were consistently observed in all subjects at all
the presented rates except the 0 wpm (resting) state. In no instance did any subject show significant activation at the 0 wpm rate. Figure 1 shows statistical probability maps for a typical subject at the four presentation rates, thresholded at a t-value of 1.83 (P < 0.05) and color coded for values above that. The images are averages of three repeats of each task as described earlier. In Fig. 1A, three distinct ROIs were observed bilaterally and have been so marked with numerals in each hemisphere on the 90 wpm overlaid image. ROIs 1 and 2 are primary auditory cortex activations mapping to lateral and medial TTG respectively (BA 41, Heschl’s gyrus). ROI 1 in the right hemisphere and ROI 2 in the left extend across the TTS into the STG. ROI 3 maps to the posterior and superior aspect of the STG in both the hemispheres (BA 22 and 42). In the left hemisphere, this represents an activation of Wernicke’s area. The activations followed sulcal folds from the lower to the upper slice and were seen again in a slice 7 mm superior to the slice shown.

To investigate the reproducibility of these results, the subject was studied twice. Figure 1B shows the observed data for a study performed 5 wk earlier. Here, data were acquired in two axial-oblique slices separated by 7 mm (as in the later study), but these slices were anatomically approximately midway between those of the subsequent study. To compare similar regions, the data from the two slices of the earlier study were averaged together. Figure 1B shows the data to be reproducible at every rate despite the slight difference in anatomy. Although the overall level of activity and the signal-to-noise is slightly different between the two sessions, the location of activations are highly similar as are the rate dependence and laterality effects.

Figure 1C shows the time course of the signal change (percent $\Delta S/S_o$) of those voxels in the primary auditory cortex with $P < 0.05$ in the left hemisphere (ROI 1 + ROI 2 in Fig. 1A). The percentage signal change (percent $\Delta S/S_o$) of the 18 voxels comprising this ROI has been plotted as a function of the acquired image number. The horizontal bar on the abscissa represents the active task period (images 5–12; stimulus started immediately after image 4 was acquired). Signal intensity started to rise after the onset of the task (image 4) and had reached its mean activated value 3.75 s later, when the subsequent image was acquired. At the 90 wpm rate, the maximum signal increase was 5.6% of the pretask baseline while the mean signal change for images 6–12 was 4.6%. Signal decay followed cessation of the task, with a time lag as shown of 11.3 s to reach the baseline value. At 90 wpm, the mean $\Delta S/S_o$ of activated voxels of all the subjects studied was 6.4%. The average delay for the signal rise to peak after commencement of the task was 3.8 s, with a much slower time lag before return to baseline of 10.5 s.

Rate dependence of the activations

Figure 1A shows a comparison of the activation profiles at different word presentation rates for a typical subject. At 10 wpm, ROIs 1 and 2 are activated weakly. There are increases in their activated volumes and $\Delta S/S_o$ values as the rate increases to 50 wpm and to 90 wpm. At 130 wpm, the activated volumes and $\Delta S/S_o$ values fall. ROI 3 also showed a similar rate dependent behavior of the activated volume and $\Delta S/S_o$ even though there was a large region activated at 10 wpm. This subject (female) demonstrated an unusually strong left dominance in both the primary and the secondary regions at all rates. This was observed to a lesser extent in the other subjects (see later).

Figure 2 is the average histogram of all six subjects plotting the behavior of voxels in the auditory cortex as defined from surface anatomy. Included are all voxels in bilateral TTG, STG, and STS. In each subject, the size of the anatomic region in which the voxels were counted was kept constant across all rates. This allows a quantitative comparison of the overall distribution of t-values. At 0 wpm, we observed the expected narrow Gaussian distribution of voxels centered at 0. At 50 wpm, there was a small shift of the distribution toward both higher positive and negative t-values, with a greater shift toward positive values that corresponds to an increase in $S_o$. The shift was more apparent in the 90 wpm curve. At 130 wpm, the distribution was approximately midway between 50 and 90 wpm. Note that the volume of activation is greater at 90 wpm than at 50 wpm independent of the t-threshold used to define activation. The voxels that appear to be activated at 0 wpm represent the statistically expected false positives at a $P < 0.05$.

Figure 3 represents the integrated signal change observed in bilateral primary and posterior superior secondary auditory cortices within the ROIs averaged across all six subjects. Here, the average percent signal change in the ROIs during activation at each rate has been multiplied by the number of activated voxels ($P < 0.05$) to provide a quantitative measure of the rate and laterality effects. By keeping the size of the region from where we assessed the fMRI signal response constant, our attempt was to quantify a dependent measure most comparable to the PET measurements of Price et al. (1992). The average volume of each of our four segregated regions was 14 ± 4 ml and the Talairach coordinates were $x = \pm 47 \pm 6$, $y = \pm 23 \pm 4$, and $z = 12 \pm 3$ mm for bilateral primary auditory cortices and $x = \pm 60 \pm 9$, $y = \pm 33 \pm 5$, and $z = 14 \pm 4$ mm for bilateral posterior superior secondary auditory cortices. Note that our activated ROIs for each subject were entirely contained within the 14 ± 4 ml defined regions and were highly similar in location and size to the approximately 15-ml voxels of Price et al. For the integrated signal thus defined, we again observed a rise in the signal with increasing rate of word presentation from 10 to 90 wpm in each of the four cortical regions investigated.

The dependence of the activations on different rates of word presentation was assessed for statistical significance by subjecting the data to an analysis of variance (ANOVA) with three within subject factors: rate (10, 50, 90, and 130 wpm); hemisphere (left vs. right); and region (primary vs. secondary). The 0 wpm rate was not included in the ANOVA since it had been established that, as expected, the fMRI data clearly showed no activation at the 0 wpm (resting) state. Instead, we were interested in using the ANOVA’s to detect higher order effects. The dependent measure used was the integrated signal change observed in the ROIs.

A Greenhouse-Geisser correction has been applied to all values to correct for repeated measures. The initial overall ANOVAs revealed significant main effects of rate.
AUDITORY CORTEX RESPONSE TO WORD PRESENTATION

FIG. 2. Histogram of voxels in bilateral primary and posterior superior secondary auditory cortices (average of 6 subjects). Vertical lines represent t = 0 and t = 1.83.

\[ F(3,15) = 15.120, P < 0.002 \] and hemisphere \[ F(1,5) = 6.643, P < 0.050 \] whereas the effect of region was not significant \[ F(1,5) = 1.858, P < 0.231 \]. None of the interactions between factors were significant.

To explicate the main effect of rate, we did pairwise two-tailed t-test comparisons among the rates collapsed across regions and hemispheres. The rate dependent increase was a trend when going from 10 to 50 wpm \( P < 0.08 \) but was significant from 10 to 90 or 130 wpm \( P < 0.0009, P < 0.009, \) respectively. The increase was significant from 50 to 90 wpm \( P < 0.006 \) and was a trend from 50 to 130 wpm \( P < 0.06 \). The decrease from 90 to 130 wpm was also a trend \( P < 0.06 \). The main effect of hemisphere was due to consistently greater activation in the left than in the right hemisphere. As stated, we were interested in testing the hypothesis that the signal has a different rate dependence in left posterior superior secondary auditory cortex as compared to the right posterior superior secondary auditory cortex and bilateral primary auditory cortices. Although the initial ANOVA provided no evidence of such a difference, we performed another series of ANOVAs using only two within subject factors: rate (10, 50, 90, and 130) and cortex (left primary, left secondary, right primary, and right secondary). The effect of cortex was not significant \[ F(3,15) = 2.454, P < 0.147 \], but there was a trend towards an interaction between rate and cortex \[ F(9,45) = 2.696, P < 0.079 \]. An inspection of the bar graphs in Fig. 3 indicates that while the rate dependence of the activation from 10 to 90 wpm is similar in the four regions, the behavior at 130 wpm is not. Indeed, when we performed ANOVA’s again without the 130 wpm data, the rate-cortex interaction became insignificant confirming that the rate dependent increase in activation from 10 to 90 wpm is highly similar in all four regions. The trend for a difference among regions when 130 wpm data are included comes from a significant decrease in activation from 90 to 130 wpm in bilateral primary and right posterior secondary regions, and a nonsignificant increase in the left posterior secondary region.

We also wanted to assess if ANOVAs conducted using signal intensity and activated volume would yield similar results. This was tested by deconvoluting the integrated signal to yield the signal intensity and the activated volume at \( t > 1.83 (P < 0.05) \). These parameters were then used as dependent measures for the new ANOVAs. Mirroring the conclusions of the data presented earlier, the ANOVAs for signal intensity and activated volume showed that there was no significant effect of either a rate-cortex interaction \( P < 0.21 \) and \( P < 0.18, \) respectively) or of cortex \( P < 0.56 \) and \( P < 0.07, \) respectively). The trend seen in the cortex ANOVA became insignificant when the 130 wpm data were not included in the analysis. However, the primary effect of rate remained highly significant \( P < 0.04 \) and \( P < 0.002, \) respectively) regardless of the data analysis strategy employed.

We accounted for intersubject differences in anatomy and brain orientation by segregating the ROIs based on the subject’s brain. In the data presented, we have assumed the boundary between secondary and primary regions to be posterior and superior to the TTS on the

FIG. 3. Rate dependent increase in integrated signal at \( P < 0.05 \) in bilateral primary and posterior superior secondary auditory cortices. Integrated signal is defined as volume of activation multiplied by average \( \Delta S/S \), and is a measure of total activity. Anatomic definition of the areas was performed using T1 weighted scout images.

A: patterns of activation in an axial-oblique image of 1 subject. Color overlays represent t-test comparisons between pretask and task periods where performed task was passive listening to words presented binaurally at 0, 10, 50, 90, and 130 wpm/word per minute (wpm). T-images have been thresholded at a value of 1.83 \( (P < 0.05) \). In this subject, no activation was seen at 0 wpm, and that image has therefore not been presented. Same color scale has been applied to all the rates with hotter colors representing increasing t-values in range 1.83–7.00. Three different regions of interest (ROIs) in left hemisphere have been labelled in 90 wpm image. ROI 1 = lateral transverse temporal gyrus (TTG); ROI 2 = medial TTG; ROI 3 = posterior superior superior temporal gyrus (STG; see Fig. 2 and text for details). B: observed activations at 10, 50, 90, and 130 wpm in same subject shown in A but from data acquired 5 wk earlier \( (t > 1.83, P < 0.05) \). C: time course of the primary auditory cortex activation in left hemisphere of inferior slice (ROI 1 + ROI 2 in A, 90 wpm). Percentage \( \Delta S/S \), of 18 pixels comprising this ROI has been plotted as a function of acquired image number. Horizontal bar on abscissa represents active task period (images 5–12).
STG. We considered the possibility that there may be some small error in the demarcation of the boundary between secondary and primary regions. Therefore we also looked at the signal in activated voxels only in the STS and in the very posterior and superior regions of the STG. Again, there were clear increases in activation as the stimulus rate increased from 10 to 50–90 wpm.

In another effort to compare our data with those of Price et al. (1992), we searched for individual voxels in each subject that had a t-value of $t = 1.83$ ($P < 0.05$) at the 10 wpm rate but then showed no significant change in t-value with increasing rate. An insignificantly small number of pixels were selected by our algorithm ($P < 0.05$). The process was repeated after smoothing the data with a 20-mm Gaussian sphere, but we still saw no regions satisfying this condition. However, even after smoothing, a search for voxels showing a t-value rising with rate consistently resulted in at least a few voxels in the ROIs (data not shown).

Finally, we recognized that an ANOVA that did include the 0 wpm data for two within subject factors [rate (0, 10, 50, and 90 wpm) and cortex (left primary, left secondary, right primary and right secondary)] and then tested for a rate-Cortex interaction would also be a close analogue of the strategy of Price et al. (1992). We performed this ANOVA and again found the effects of cortex and the rate-cortex interaction to be nonsignificant ($P < 0.34$ and $P < 0.58$, respectively), whereas the primary effect of rate remained highly significant ($P < 0.001$).

**DISCUSSION**

Although fMRI previously has been used successfully to localize brain activations during visual, sensory, and motor tasks (Blamire et al. 1992; Kim et al. 1993; McCarthy et al. 1993), the investigation of auditory processing presents a special challenge. Background noise generated by the gradients required for image formation potentially raises the baseline typically used to calculate the difference in the NMR signal in the pretask and the task states. However, our results show that fMRI clearly demonstrates the ability to detect distinct brain activity in primary and secondary auditory cortices, in agreement with the recent results of Binder et al. (1994a,b).

During the presentation of pure English nouns, we observed localized activations in the TTG and the TTS. These are consistent with the traditional implication of these areas in the early processing of auditory input. In addition, we observed bilateral foci of activations posterior and superior to the primary auditory cortices that we identified as secondary activations. The focus of activation in the left posterior secondary region lies in Wernicke’s area. These regions of activation are similar to those reported in previous PET and fMRI studies of auditory processing (Binder et al. 1994a,b; Petersen et al. 1988; Zatorre et al. 1992).

We observe a rise in the integrated signal with increasing rates of word presentation. In the activated ROIs, we see a continuing rise in $\Delta S/S$, with rate accompanied by a significant rise in the number of activated voxels in bilateral primary and posterior secondary auditory cortices. Precise definition of the volume of activation during a particular task condition is difficult because the volume depends on the somewhat arbitrary selection of a threshold for the t-value. However, as may be seen in Fig. 2, activation is more extensive at 90 wpm than it is at 50 wpm with any selected t threshold.

The statistically significant increase in integrated signal associated with increasing stimulus presentation rate is consistent with previous PET and fMRI studies of the auditory cortices (Binder et al. 1994a,b; Price et al. 1992). In addition, we have established an upper limit of the rate dependent increase with monosyllabic English noun stimulus by showing no significant increase in signal from 90 to 130 wpm, and, in fact, showing a trend toward a decrease in signal.

As in the earlier studies, with the experimental protocol used, we are unable to distinguish between the effects of changes in stimulus rate and changes in stimulus duration. As described in METHODS, increasing the word presentation rate (10, 50, 90, 130 wpm) increased the stimulus duration in a nonlinear manner (7, 34, 52, 55 s). To differentiate the effects of these different aspects of stimulation, it will be necessary for future studies to keep the duration constant while varying the rate and vice versa. But in the context of this study, the acoustic parameters of our stimulation protocol should be highly similar to those of Price et al. (1992), as also should be the physiological response.

Price et al. (1992) reported a linear rise in signal from 0 to 90 wpm in all the regions examined except Wernicke’s area. Although our data agree with the PET results in bilateral primary and right secondary auditory cortices in so much as they show a steady rate dependent rise in signal, they do not show the sharp stepwise rise in Wernicke’s area. In Wernicke’s area, our results are indistinguishable from the other regions, in contrast to the PET results. Indeed, rate dependence of left posterior superior secondary cortex persisted even when we examined only the most posterior regions, thereby eliminating any possible mixing of primary auditory cortex voxels with Wernicke’s area.

We considered the possibility that because we had acquired data in only two or four slices, we may have missed the regions investigated by Price et al. (1992). We assessed this possibility by calculating Talairach-Tournoux coordinates for each of our four segregated regions. Of the six distinct regions identified by Price et al. (1992), the bilateral primary auditory activations were centered at $x = \pm 50$, $y = -22$, and $z = 12$ mm. The average Talairach coordinates for our primary activations were $x = \pm 47 \pm 6$, $y = -23 \pm 4$, and $z = 12 \pm 5$ mm. The posterior superior temporal gyrus activations in Price et al. (1992) were centered at $x = \pm 58$, $y = -34$, and $z = 12$ mm. The average Talairach coordinates.

---

1. The data processing of Price et al. (1992) involved intersubject averaging of data, which included a 20-mm Gaussian smoothing filter to account for a scattering of anatomic structures between subjects. After creating t-statistical parametric maps of their condition means, they conducted a search for the voxels that showed a linear rise in signal with the presentation rate of heard words and found those to be located bilaterally in primary auditory cortices and medial STG and unilaterally in the right posterior superior STG. A search for voxels where the difference between rest and 10 wpm was greater than the average difference between each word presentation rate only showed voxels in the left superior posterior STG (Wernicke’s Area).

2. Our search algorithm 1) accepted the t-statistical map ($t > 1.83$, $P < 0.05$) for the 10 wpm stimulus as input; 2) performed a linear regression through the t-images of 10, 50, 90, and 130 wpm for those voxels present in 1); and 3) returned the slope and the correlation coefficient.
for these regions in our study were $\pm 60 \pm 9$, $y = -33 \pm 5$, and $z = 14 \pm 4$ mm.

In addition, from the figures of Price et al. (1992), we calculated an approximate lower estimate for the volume of activation of their ROIs to be $2 \times 3 \times 2.5 = 15$ ml. On inspection of the Talairach coordinates of the ROIs that localized to the primary and posterior secondary auditory cortices of our six subjects, we found that the individual coordinates were contained within the above PET volume in every instance. We therefore can conclude that we investigated, and have reported the results for, anatomic regions highly similar to those identified in the PET study.

Finally, on following the data analysis strategy of Price et al. (1992) and searching our data for voxels activated at 10 wpm but showing no further increase in activation at higher stimulation rates, we failed to find any. Of course, the Price et al. study directly measured rCBF differences while the signal in fMRI is believed to be an indirect, albeit close, correlate of flow changes. Previous studies have established the technique as a reliable indicator of neuronal activation in human brain primary, sensory, and cognitive areas, so it is unlikely that the differences in our results stem from inherent methodological incompatibilities (for recent reviews, see Aine 1995; Shulman et al. 1993). Moreover, we find ourselves in excellent agreement with the localization of activation as well the rate dependence of the activation from 10 to 90 wpm in bilateral primary as well as right posterior secondary cortices.

In conclusion, we have shown that in the human brain auditory cortex, increasing the processing load significantly increases the integrated signal, the number of activated voxels and the signal intensity in a fixed set of voxels. These increases occur in bilateral primary and posterior superior secondary auditory cortices and there is no detectable difference in the activation profiles of various regions within the auditory cortex.

The authors thank Dr. Fahmeeed Hyder for comments and Dr. Alex Stevens for assistance with the ANOVAs. This work was supported by National Institute of Diabetes and Digestive and Kidney Diseases Grant DK-34576. Address reprint requests to A. Dhankhar.

Received 11 December 1995; accepted in final form 3 September 1996.

REFERENCES


