◆ Human Brain Mapping 7:213–223(1999) ◆

"Sparse" Temporal Sampling in Auditory fMRI

Deborah A. Hall,[1]* Mark P. Haggard,[1] Michael A. Akeroyd,[1] Alan R. Palmer,[1] A. Quentin Summerfield,[1] Michael R. Elliott,[2] Elaine M. Gurney,[1] and Richard W. Bowtell[2]

[1]*MRC Institute of Hearing Research, University Park, Nottingham, UK*
[2]*Magnetic Resonance Centre, School of Physics and Astronomy, University of Nottingham, Nottingham, UK*

Abstract: The use of functional magnetic resonance imaging (fMRI) to explore central auditory function may be compromised by the intense bursts of stray acoustic noise produced by the scanner whenever the magnetic resonance signal is read out. We present results evaluating the use of one method to reduce the effect of the scanner noise: "sparse" temporal sampling. Using this technique, single volumes of brain images are acquired at the end of stimulus and baseline conditions. To optimize detection of the activation, images are taken near to the maxima and minima of the hemodynamic response during the experimental cycle. Thus, the effective auditory stimulus for the activation is not masked by the scanner noise.

In experiment 1, the course of the hemodynamic response to auditory stimulation was mapped during continuous task performance. The mean peak of the response was at 10.5 sec after stimulus onset, with little further change until stimulus offset. In experiment 2, sparse imaging was used to acquire activation images. Despite the fewer samples with sparse imaging, this method successfully delimited broadly the same regions of activation as conventional continuous imaging. However, the mean percentage MR signal change within the region of interest was greater using sparse imaging. Auditory experiments that use continuous imaging methods may measure activation that is a result of an interaction between the stimulus and task factors (e.g., attentive effort) induced by the intense background noise. We suggest that sparse imaging is advantageous in auditory experiments as it ensures that the obtained activation depends on the stimulus alone. *Hum. Brain Mapp. 7:213–223, 1999.*

Key words: sparse imaging; scanner noise interference; MR signal-to-noise ratio

INTRODUCTION

Functional magnetic resonance imaging (fMRI) is receiving increased attention as a method to investigate brain activation by auditory stimulation [e.g., Binder et al., 1994a,b; Millen et al., 1995; Woodruff et al., 1996]. However, the application of functional neuroimaging to auditory research is more problematic than for other sensory modalities for several reasons: 1) the incomplete knowledge of the function of the auditory cortex, 2) the small size and nonsuperficial location of the human auditory cortex, 3) the difficulty of delivering high-fidelity calibrated stimulation in the high magnetic fields, and 4) the intense masking noise generated by the MR scanner. This paper addresses the last issue, but starts with a brief introduction to the first three.

Grant sponsor: Medical Research Council; Grant number: G9302591.

*Correspondence to: Deborah A. Hall, MRC Institute of Hearing Research, University Park, Nottingham, NG7 2RD, UK. E-mail: debbie@ihr.mrc.ac.uk

Received for publication; 26 May 1998; accepted 28 October 1998

UNDERSPECIFIED NEUROPHYSIOLOGY

There appear to be considerable differences between the auditory cortex and that of other sensory modalities. Despite a wealth of basic studies of the auditory cortex of various species [Aitkin, 1990], it has been difficult to establish such clear-cut organizational principles or functions as found in the visual and somatosensory cortex [Zeki, 1993]. Certainly, a number of auditory cortical areas, including the primary auditory cortex, are represented topographically (cochleotopic organization) and this organization has been identified in the primary auditory cortex of humans using positron emission tomography (PET) [Lauter et al., 1985] and fMRI [Talavage et al., 1997; Wessinger et al., 1997]. Further, the primary auditory cortex of the cat and other animals seems to be topographically organized with respect to other stimulus features, such as threshold and sharpness of tuning [Shamma et al., 1993; Schreiner and Mendelson, 1990; Sutter and Schreiner, 1991; Kowalski et al., 1995]. However, this organization has yet to be studied in humans. It is clear that we do not have the same depth of understanding of the function, organization, or interconnections of the various cortical areas subserving audition as exists for vision. Consequently, we lack a strong basis for hypothesis generation or detailed interpretation of findings in auditory neuroimaging.

ANATOMICAL CONSTRAINTS

The primary auditory cortex is relatively small (between 1–4 cm^3 in each hemisphere) and its position and extent within Heschl's gyrus are both asymmetric and highly variable across individuals [Penhune et al., 1996]. This makes the accurate identification of those voxels lying within the primary auditory cortex difficult. Furthermore, the auditory cortex is located relatively close to sinuses, where interfaces between air and bone occur. At such interfaces, differences in magnetic susceptibility can cause image artifacts such as spatial compression at the edges of the grey matter tissue and loss of signal intensity due to the dephasing of spins within individual voxels [Henkelman and Bronskill, 1987]. Susceptibility artifacts such as these are increasingly evident at higher field strengths.

DIFFICULTIES IN HIGH-QUALITY SOUND DELIVERY

The high magnetic fields dictate that no magnetic materials may be used in the vicinity of the scanner: even nonmetallic materials may not be completely magnetically inert in this situation. This poses a problem for presenting subjects with low-distortion auditory stimuli across a wide frequency range. Many current systems utilize loudspeakers from which sound is delivered through tubes inserted into subjects' ears through a protective ear-defender. The tubing affects both the phase and amplitude of the different frequency components of a stimulus and makes it unfeasible to implement active noise cancellation techniques. We have developed a calibrated, high-fidelity headphone sound system for use in MR scanners [Bullock et al., 1998] that permits controlled stimulation and offers the potential for a level of active noise cancellation that is likely to be useful in a fully-engineered system.

STRAY ACOUSTIC SCANNER NOISE

The intense stray masking noise resulting from mechanical forces created by the switching of the gradient coils every time the MR signal is read out creates a severe problem for auditory fMRI studies, particularly since it is necessary to acquire large numbers of images during a functional imaging session. It is this problem that we address in this paper.

The noise of the scanner has the character of a complex tone. Its spectrum varies according to the type of pulse sequence implemented, but it is always characterized by high-intensity (110–140 dB sound pressure level (SPL)) spectral peaks at the switching periodicity and its (odd) harmonics, mostly within the frequency range of 0–3 kHz [Hedeen and Edelstein, 1997; Ravicz et al., 1997]. Unfortunately, this frequency range is crucial for speech intelligibility. Typical functional imaging paradigms measuring state-related responses consist of repeated cycles of stimulation and baseline epochs, throughout which an MR signal is measured at regular intervals. This experimental protocol produces repeated bursts of scanner noise at a rate that broadly reflects the time between slice acquisitions. The auditory system is therefore subject to continuous quasi-tonal stimulation, resulting in an elevated baseline level of activation. Stimulus-induced activations tend to be on the order of 2–6% from baseline. Thus higher levels of baseline activation, caused by the ambient noise, are likely to make the experimentally-induced auditory activation more difficult to detect statistically. Phenomenologically the scanner noise makes the stimulus more difficult to hear. Indeed, several studies have reported an enhanced activation signal (i.e., the difference between stimulation and baseline conditions) in the auditory cortex when the amount of prior gradient noise is reduced, indicating that the scanner noise does mask the detection of auditory activation [Bandettini

et al., 1998; Shah et al., 1997]. With an active auditory task that requires discrimination between stimuli, rather than passive listening, the scanner noise may even affect the auditory processes themselves by increasing the attentional load.

The concern for undetermined and spurious interactions between brain activity and additional acoustic noise parallels similar unease about the assumption of ''pure insertion'' in cognitive subtraction paradigms. ''Pure insertion'' states that an additional cognitive component can be added to a task with no effect on the nature of processing within the previous cognitive components [Price and Friston, 1997]. However, if the assumption of pure insertion is violated, then the obtained activation reflects not only the added task component, but also the interaction between added and existing components. Under continuous scanning conditions, in auditory fMRI studies, the activation detected is assumed to be caused by the stimulus or task alone. We suggest instead that in some experimental paradigms, the activation may reflect not only stimulus-induced activation, but also its interaction with background acoustical noise. While steps can be taken to reduce the intensity of scanner noise, such as improved sound insulation at the ear or active noise cancellation, it cannot be eliminated completely.

Scanner acoustic noise particularly stands to limit what might be learned from any detailed variation of stimulus parameters. Paradigm optimization must therefore aim to reduce scanner acoustic noise, particularly when subtle auditory responses are to be measured. The necessity for noise reduction will be partly determined by an experiment's rationale. For example, noise interference may be an acceptable tradeoff when examining ''oddball'' responses to single novel or rare events, for which the preferred imaging protocol would be event-related fMRI.

The auditory response to scanner noise can be characterized as interfering with the stimulus-induced auditory response along two different temporal scales. First, the scanner noise generated by the acquisition of one slice in the volume may induce activation in an imaging slice which covers the auditory cortex and is acquired later in the same volume. Second, scanner noise may induce auditory activation that extends across time to subsequent volumes. Different continuous imaging sequences can reduce some aspects of the interference from scanner noise. For example, Talavage et al. [1998a] found that a clustered-volume acquisition sequence (in which all images are acquired at one end of the time to repeat (TR) period) could reduce the impact of scanner noise on later images of the same volume, compared with a distributed volume acquisition sequence (in which images are acquired equally throughout the TR period). The clustered-volume acquisition sequence was found to maximally reduce interference across images of the same volume when the duration of the scanner noise was 2 sec or less [Talavage et al., 1998b]. However, this sequence would do little to reduce the impact of scanner noise from one volume to the next because the effective TR remains identical to that in distributed-volume acquisition.

In this paper, we address the design of efficient paradigms for use when it is desirable to reduce the impact of acoustic scanner noise on auditory activation. We report one method which we term ''sparse'' temporal sampling. Sparse imaging uses a clustered-volume acquisition sequence to reduce *intra*volume noise interference. To reduce *inter*volume noise interference, we reduce the rate of the bursts of scanner noise by increasing the interval between each set of data acquisitions (i.e., increasing TR). This ensures that the measured activity in the auditory cortex is uncontaminated by its responses to the preceding burst of scanner noise. Sparse temporal sampling is characterized by the acquisition of only *one* volume during each epoch. We therefore seek to acquire images near to the maxima and minima of the mean hemodynamic response since imaging the auditory cortex at these two time points will enhance signal detection. The nature of the hemodynamic response is determined in part by local cerebral vasculature and may therefore vary across both cortical regions and individual subjects. In the primary visual and motor cortex, the time to reach 90% of the peak change in the MR response occurs approximately 5–8 sec from stimulus onset [Blamire et al., 1992; Kwong et al., 1992]. In contrast, there are no direct estimates of the hemodynamic delay in the auditory cortex during periods of auditory stimulation. There are several reasons for obtaining an accurate characterization of the hemodynamic response to auditory stimulation. First, the determination of the mean hemodynamic delay in the auditory cortex is necessary to optimize the synchronization of image acquisition with stimulus presentation when using sparse imaging. Second, if the mean hemodynamic delay differs greatly across individuals, it may be important to adapt the temporal aspects of the sparse imaging paradigm to suit individual hemodynamic characteristics.

In experiment 1, we measured hemodynamic response by collecting samples of images at short intervals during a simple auditory task which involved passive listening to continuous speech. In experiment 2, we evaluated the potential of the sparse temporal sampling technique as a method for auditory fMRI.

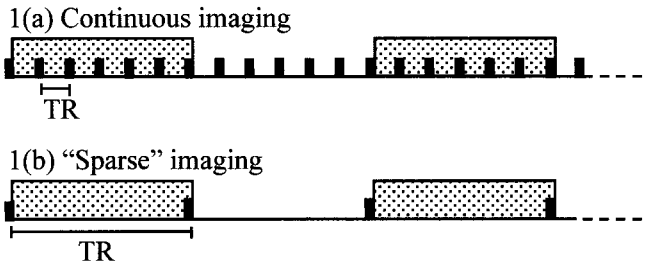

Figure 1.
A schematic representation of the two imaging paradigms used. Diagram illustrates different rates of image acquisition and their synchronization with stimulus onset. Wide horizontal bars indicate periods of stimulation, and short vertical lines indicate acquisition of a single volume of data. In each experiment, epochs were 14 sec in duration. In continuous imaging, the interscan interval (TR) was 2.33 sec, and in sparse imaging it was 14 sec.

METHODS

Both studies were performed on a dedicated echo-planar imaging (EPI) 3 Tesla scanner with purpose-built head gradient coils and a birdcage radio frequency (RF) coil [Bowtell et al., 1994]. For both studies, the image matrix contained 128 × 128 elements, and the resolution of single voxels was 3 × 3 × 8 mm. An modulus blipped echo-planar single-pulse technique (MBEST) sequence using clustered-volume acquisition permitted the acquisition of a set of eight contiguous coronal slices in 536 msec. Slices were chosen to cover the primary and association auditory regions of the temporal lobes.

The stimuli were presented using a specially engineered sound system which delivers sounds using electrostatic headphones combined with standard industrial ear defenders [Bullock et al., 1998]. Auditory stimulation was prerecorded continuous speech taken from a "talking book" that had been segmented at 14-sec intervals. In both tasks, subjects were instructed to lie motionless with eyes closed and listen to the spoken story. Subjects heard alternating intervals of 14 sec of speech and 14 sec of silence, presented for either 32 or 40 repetitions. Figure 1 shows a schematic diagram of the two different data sampling techniques.

All subjects were right-handed adults. They had no history of neurological or auditory impairment, and were not on any medication. Prior to the imaging session, the hearing sensitivity of subjects was measured using pure-tone audiometry. The hearing thresholds of all subjects fell within the normal range (<20 dB hearing level) at octave frequencies between 500–8,000 Hz, inclusive.

Experiment 1

Seven normal subjects, aged 22–46, participated in experiment 1. A set of images was acquired every 2.33 sec (TR = 2,330 msec; TE = 27 msec), sufficiently frequently to map out the rise and fall of the hemodynamic response. This method is denoted "continuous imaging" as it involves frequent sampling of the MR signal. Five sets of "dummy" images were taken before the onset of the first speech epoch to allow time for the longitudinal magnetization to reach a steady state. These data were discarded prior to statistical analysis. Thus, in total we acquired 384 or 480 (32 or 40 repetitions, respectively) sets of functional images per subject. The timing of the first set of images corresponded to the exact onset of the initial speech epoch (see Fig. 1a).

Experiment 2

Six normal subjects, aged 19–35, participated in experiment 2. For half the subjects, TE = 27 msec, and for the other half TE = 36 msec. Only 3 of these subjects had participated in experiment 1 (i.e., those for whom data were obtained using TE = 27 msec). In all other respects the imaging parameters remained identical. A volume of images was acquired once every 14 sec (TR = 14,000 msec). The time interval between data acquisitions was based on previous work in mapping the time course of visual and motor activation [e.g., Blamire et al., 1992; Kwong et al., 1992] and was sufficiently long to allow the hemodynamic response to reach a plateau. Images were acquired at the end of each speech and silent epoch (see Fig. 1b), and 64 volumes of images were acquired in total.

For an accurate comparison between continuous and sparse imaging, it was necessary to acquire data using both techniques within the same experimental session. Continuous imaging used the same imaging sequence as in experiment 1 (i.e., an image volume was acquired every 2.33 sec). The same-session control helped minimize within-subject differences in head position across the two experiments as well as differences in signal contrast to noise ratio due to quality of shimming and RF coil tuning.

RESULTS

The data from experiments 1 and 2 were analyzed according to the general linear model using SPM96 [Friston et al., 1994]. For individual subjects, data were realigned to the average of the images in the experimen-

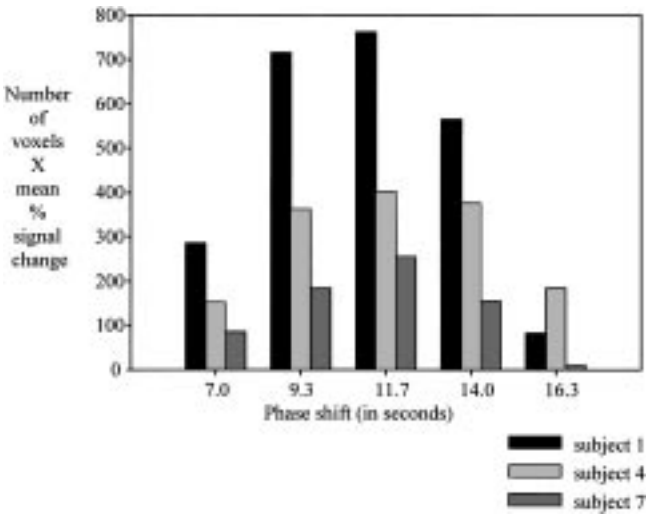

Figure 2.

Preliminary analysis showing the distribution of activation in the auditory cortices across time in a subset of subjects (N = 3) in experiment 1. Amount of activation is plotted against contrasts, using sine-wave basis functions, each shifted in phase by one scan (2.33 sec). Each reference waveform correlates most strongly with those voxels showing a sinusoidal change in activation that peaks at the point corresponding to the phase shift. Thus, for this group, maximal activation appears to occur approximately 11.65 sec following onset of activation.

tal sequence and were corrected for three-dimensional head movement. Realigned images were then spatially smoothed with a Gaussian kernel (FWHM 5.5 mm). Low-frequency artifacts, corresponding to aliased respiratory and cardiac effects and other cyclical variations in signal intensity, were removed by high-pass filtering the time series, using cosine basis functions up to a maximum of one cycle per minute. Image data were not smoothed in the time domain.

Experiment 1

In experiment 1, we modeled a difference in the MR response between silence and listening to speech by using a sine-wave basis function at the fundamental frequency of the experimental cycle. The general linear model incorporated five such sine waves that were incrementally shifted in phase by one scan (i.e., 2.33 sec) in order to identify those voxels whose signal changed across different hemodynamic phases. We used phase shifts from 3–7 scans (i.e., 7–16 sec). Preliminary analysis of the subset of the data (N = 3) showed that this range of reference functions appropriately detected the time-course of activation changes in the auditory cortex (Fig. 2). When the sine-wave peak was located at the third scan (i.e., 6.99 sec) or at the seventh scan (i.e., 16.31 sec) from stimulus onset, there was very little activation. For these 3 subjects, activation was greatest when the sine wave peak was located at scan 5 (i.e., 11.65 sec) from stimulus onset.

Changes in the corrected MR signal within each voxel that followed these five reference waveforms were identified for all subjects using the F statistic. The spatial distribution of statistical significances was expressed as an SPM{F}. map thresholded at $P < .0001$, uncorrected for multiple comparisons. For all subjects, this analysis successfully delimited bilateral regions of activation in the auditory cortex, including the primary auditory cortex (Heschl's gyrus) and superior temporal gyrus.

The second and most important part of the analysis calculated the time course of the MR signal within the identified areas of auditory cortex. The maxima and minima of the MR signal changes provided a measurement of the lag of the hemodynamic response in this region of cortex. Within a region of activation, the time course of the corrected MR signal was extracted for each voxel. A region was defined as a cluster of adjacent voxels that were expressed in the SPM{F} map at the $P < .0001$ threshold of significance. Within a region of interest, the data were averaged for each subject across all active voxels and across experimental cycles. Signal intensity changes were standardized by calculating the mean percentage change relative to the baseline within each region. Figure 3 shows the mean time-course of the MR signal, within the region of interest, for each of the 7 subjects, fitted by a smoothed spline curve. The time-course data were averaged across the group, and a smoothed spline curve was fitted to the data (Fig. 4). The time to reach the absolute peak in activation change was 10.5 sec from stimulus onset. The standard deviation across the group in terms of the hemodynamic peak was 1.44 sec. The preliminary analysis of 3 subjects (shown in Fig. 2), which indicated that the peak value was around 11.65 sec, is consistent with the calculation based on the group analysis, since it falls within one standard deviation of 10.5 sec.

The hemodynamic lag was also measured at the point on the curve at which the data reached within 10% of the absolute maxima and minima. The 10% criterion has been widely used to estimate the latency of the hemodynamic response in the primary visual and motor cortex [e.g., Bandettini et al., 1993; Blamire et al., 1992; Kwong et al., 1992]. This method yielded a value for the mean latency of the hemodynamic maxi-

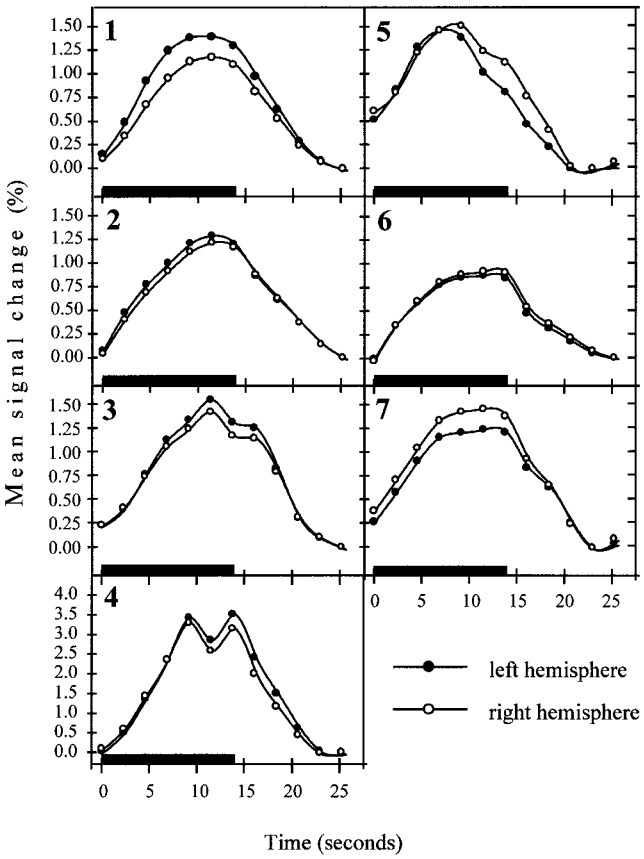

Figure 3.
Individual hemodynamic responses averaged across active voxels within primary and secondary auditory cortex, as measured using continuous imaging in experiment 1.

mum in the auditory cortex at 7.7 sec from stimulus onset and of the minimum at 8.1 sec from stimulus offset.

Under the stimulus conditions used, once a peak in the MR signal change was reached, the mean hemodynamic response plateaued out until the offset of the stimulus (Fig. 4). Pairwise analyses on the mean signal change at time points 9.3, 11.7, and 14.0 sec after stimulus onset, for individual subjects, revealed that the MR signal change did not differ significantly from the peak to the end of the epoch (9.3 sec vs. 11.7 sec, $t[13] = 0.72$, ns; 11.7 sec vs. 14 sec, $t[13] = 1.13$, ns; and 9.3 sec vs. 14 sec, $t[13] = 1.8$, ns).

The shape of the hemodynamic response appears to vary somewhat across individuals (see Fig. 3). The overall mean and SD across the group for the hemodynamic rise to 90% were 7.7 sec and 1.4 sec, respectively. The mean ranged across subjects from 4.9–9.8 sec. The differences in lag time across subjects were significant ($F[6, 7] = 27.10$, $P < .001$). This variation may be due to intersubject differences in the exact position of blood vessels across the Sylvian fissure, which will affect local hemodynamics. Local blood flow will affect voxel-based distributions of MR signal magnitude and delay that depend on the precise location of those voxels in relation to major venous drainage. However, the region of interest here was delimited on the basis of statistical significance, not on the basis of specific structural considerations. Therefore, the number and location of voxels used to compute a subject's hemodynamic response varied slightly across individuals. The intrasubject variability in terms of the shape of the hemodynamic response across hemispheres was smaller than the intersubject variability. For example, in terms of lag time, there was no within-subject difference between the left and right active auditory regions ($F[1, 12] = 0.18$, ns). It is therefore possible that much of the intersubject variation lies in global differences of the hemodynamic response itself.

Experiment 2

In experiment 2, the sparse temporal sampling technique provided image data at two time points, one in each of the stimulation and baseline epochs. The MR

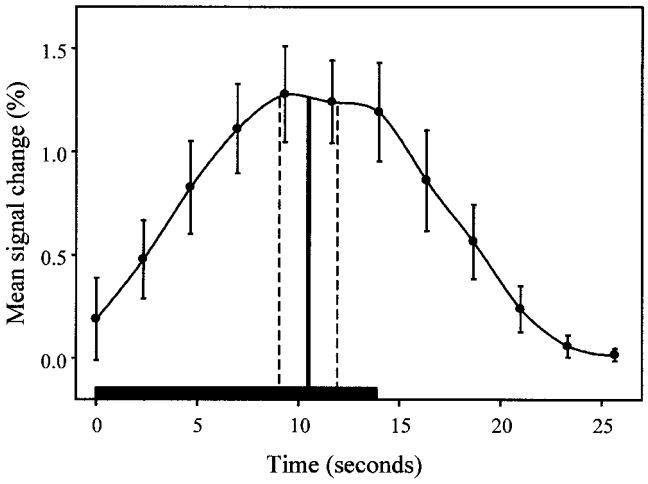

Figure 4.
Mean hemodynamic response across subjects in experiment 1 during stimulation and baseline conditions, where timing of stimulus presentation is indicated by horizontal black bar. Percentage signal change was obtained by averaging time course of those activated voxels within primary and secondary auditory cortex. Variability across subjects is indicated by error bars (±1 standard deviation from the mean). Solid vertical line indicates point along the spline curve at which hemodynamic response reaches maximum. Dotted vertical lines at ±1 standard deviation indicate distribution of the hemodynamic lag across the group.

TABLE I. Extent and mean percent signal change with continuous and sparse imaging in experiment 2

	Continuous imaging		Sparse imaging	
Subject	No. of voxels	Mean % MR signal change	No. of voxels	Mean % MR signal change
TE 26				
1	125	2.16	303	1.57
2	261	0.95	47	1.52
3	438	0.92	256	1.32
TE 36				
4	175	1.48	315	2.91
5	495	1.45	299	2.29
6	424	1.93	479	3.60

response was modeled using a simple square-wave function where the data points fell on the maxima and minima of the square-wave. The MR response to auditory activation was analyzed by subtracting the baseline from the stimulation states, thus identifying those voxels whose intensities significantly differed across the stimulation and baseline conditions. The continuous imaging data that had been acquired during the same imaging session were analyzed in the same way as in experiment 1. For both the continuous and the sparse data sets, the spatial distribution of statistical significances was expressed as an SPM{F} map thresholded at $P < .001$, uncorrected for multiple comparisons. For each subject, the same extent threshold was used across the two analyses to identify activated brain regions. As shown in Table I, there were large within-subject differences in the extent of activation across continuous and sparse imaging data sets. However, discrepancies between the directions of this difference made it nonsignificant (paired t [5] = 0.22, ns). We conclude that the two imaging techniques are generally equivalent in terms of their ability to detect clusters of voxels activated by listening to speech compared with silence. However, this result does not rule out important differences in more specific respects, such as the size of the signal change.

For each of the 6 subjects, the activation maps for the two conditions were overlaid onto an average of the individual's T2*-weighted functional scans. From these overlays, it was possible to identify the anatomical positions of areas of activation. Six functional auditory regions within the temporal lobe were identified according to Rivier and Clarke [1997]. The data set from each subject was categorized according to whether or not activation was present in each of these areas. All 6 subjects showed bilateral activation in the primary auditory cortex, the lateral and posterior auditory areas, the superior temporal gyrus (which includes Wernicke's area), and the middle temporal gyrus. Four out of the 6 subjects also showed activation in the anterior auditory area. The small number of subjects obviates reliable statistical analysis of any differences between methods. The categorical data from all 6 subjects were therefore pooled for descriptive purposes: 6 subjects × 6 areas × 2 hemispheres. There was 79% (57/72) agreement between the two imaging methods on areas that were identified as containing significant activation. In 10 of the cases that disagreed, an area where activation was detected during continuous imaging was not obtained during sparse imaging. In 5 cases the converse pattern was achieved. The differences in the *differential* activation pattern between the two methods were distributed over most of the auditory regions, including the primary auditory cortex. A possible explanation for the overall direction of the discrepancy, which led to a slight increase in the number of functional regions active using the continuous imaging method, may be the extra effort associated with the attention to the masked stimulus in continuous imaging. We conclude that there are no major overall differences between the two imaging methods in terms of the regions activated by listening to speech. Figure 5 presents the activation patterns for one subject, illustrating similar patterns of activation across the two imaging paradigms.

For each of the 6 subjects, the mean percentage signal change was calculated across all significantly active voxels within bilateral regions of interest (Table I). For the sparse imaging data, this value was derived from the difference between the corrected MR signal at the stimulus and baseline acquisitions. For the continuous imaging data, the mean percentage signal change was calculated from the difference between the MR signal at the two equivalent time points (i.e., at the transitions between stimulus and baseline epochs). Figure 5 shows that, for this particular subject at least, the peak voxel activation was higher in the continuous imaging experiment. However, when the mean signal change was calculated across regions of interest, 5 out of the 6 subjects showed less activation with continuous imaging. This difference is likely to be partly a consequence of the effect of different rates of imaging (i.e., TR) on the SNR of the image. In continuous imaging, the rate of image acquisition is such that the net magnetization has not recovered to equilibrium between excitations. This effect reduces the magnitude of the MR signal relative to that obtained using sparse imaging, and hence reduces the signal to noise ratio (SNR). The mean percentage signal change is also

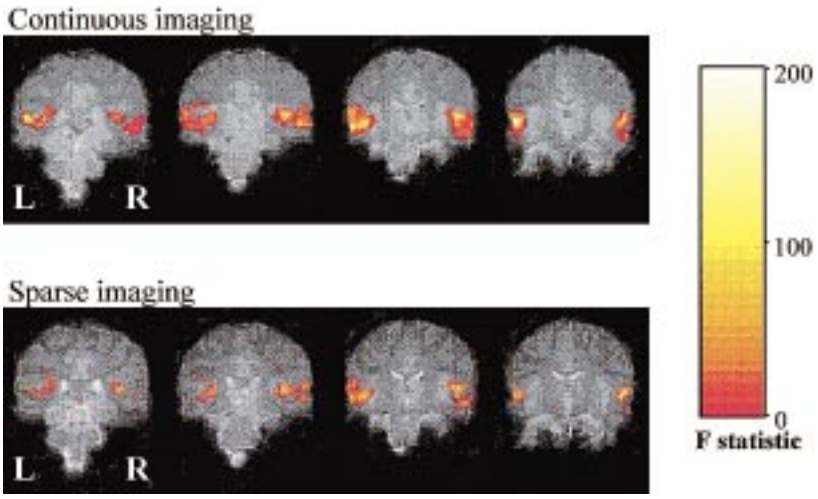

Figure 5.

Pattern of activation in a single subject who was scanned using the two different methods of temporal sampling in experiment 2. Activation maps were thresholded at $P < .001$ and were uncorrected for multiple comparisons. Regions of activation are shown whose size exceeds 100 voxels. Both techniques succeeded in detecting activation in transverse temporal gyrus, superior and middle temporal gyrus, and a small area of activation in the inferior parietal lobe.

generally greater at the longer echo time (i.e., TE), for both continuous and sparse imaging. Echo time determines the amount of T2* contrast that is seen in the image, and this can vary according to the field strength and local magnetic susceptibility. Our results suggest that at 3 T, a TE of 36 msec may be better than a TE of 27 msec for detecting small activation changes in the auditory cortex.

DISCUSSION

We have reported a ''sparse'' imaging technique, which reduces the rate of bursts of scanner acoustic noise by increasing the duration of the interscan interval. Compared with conventional ''continuous'' imaging procedures, sparse temporal sampling is equally effective at detecting auditory activation in terms of the extent and location of significantly activated voxels.

We suggest that the sparse imaging paradigm is advantageous in auditory experiments for three reasons. First, sparse imaging ensures that the activation is not a result of some interaction between the stimulus and scanner noise, either in terms of acoustic masking or increased attention and effort. In contrast, the uncontrolled and unknown effects of scanner noise during continuous imaging on the pattern of activation are a real concern for the interpretation of auditory fMRI data obtained using these methods. Indeed, some auditory fMRI experiments using continuous imaging have revealed activation of extraauditory centers that are not seen in PET when using the same stimulus paradigm (Griffiths, personal communication).

The second advantage of sparse imaging is that the MR signal-to-noise ratio is enhanced, despite fewer data samples than with conventional continuous sampling techniques and the consequent loss of statistical power. It may theoretically be possible to increase statistical power by increasing the number of volumes acquired. Such a technique could be described as a ''quasi'' sparse imaging method in which scanner noise interference is maintained at a minimum by acquiring multiple volumes at one end of the epoch. However, the enhanced statistical detection of activation has not yet been realized in practice. For example, to compare the quasi-sparse with the sparse imaging method, we ran a single-subject experiment using the same stimulus presentation as in experiment 2. In the quasi-sparse method, four eight-slice volumes were acquired in 2.1 sec (time between slices = 67msec) every 14 sec at the end of each stimulus and baseline epoch. Thus the time between the last volume of the stimulus epoch and the first volume of the baseline epoch was 11.9 sec. The irregular interscan interval reduced the magnitude of the MR signal and hence the SNR for each of the four volumes acquired sequentially during each epoch. However, the mean and the variance of the voxel intensities showed a strong correlation over time. Therefore, these predictable fluctuations can be adjusted by scaling by a constant

those voxel intensities for volumes taken at the same time points such that the variance falls within a normal distribution and the data can be subjected to statistical parametric mapping. Using this method of analysis, the quasi-sparse method was in fact significantly inferior in detecting areas of auditory activation. We suggest that the reduction in SNR across successive volumes contributes to the decrease in statistical power when those successive volumes are integrated in a single analysis.

In the sparse imaging method, the detection of the stimulus-induced signal is maximized during data acquisition in two ways: 1) by maximizing the difference between the two points on the hemodynamic response through synchronization of the data acquisition with the stimulus cycle, and 2) by maximizing the magnitude of the T2*-weighted signal as a result of the greater signal recovery that occurs between excitations. The magnitude of the MR signal decreases with a reduction in TR, particularly if TR is less than five times the T1 relaxation time for the tissue [Rinck, 1993]. A consequence of the reduction in magnitude of the MR signal is a concomitant reduction in signal-to-noise ratio (SNR) of the image during continuous imaging. Indeed, in sparse imaging, a greater mean MR signal-change within the region of interest was measured in all but one subject, compared with continuous imaging. A broad correspondence between the activation maps using continuous and sparse imaging methods was achieved, despite the reduction in the number of data acquisitions in the sparse imaging experiment. The third advantage of sparse imaging reflects the discomfort and even stress caused by the loud scanner noise during data acquisition. Where performance of nontrivial tasks is required, reducing the ambient noise level is desirable to avoid imaging any cognitive interaction effects arising from the symptoms of such stress.

Despite these three advantages, sparse imaging may not be the method of choice in every auditory fMRI experiment, e.g., if fine temporal resolution is required, as with event-related fMRI.

Hemodynamic response in auditory cortex

Optimization of the difference between the stimulation and baseline conditions can be facilitated following mapping of the hemodynamic response function within the auditory cortex. This was done in experiment 1 using a temporal resolution of 2.3 sec. Within those voxels that were significantly active during speech, the mean hemodynamic delay was 7.7 sec from the onset of stimulation to within 10% of the maximum, and was 8.1 sec from stimulus offset to within 10% of the minumum. The 10% criterion has been widely used to estimate the latency of the hemodynamic response in the primary visual and motor cortex. For example, in these cortical regions the response is approximately 5–8 sec from stimulus onset to 90% peak, and 5–9 sec from stimulus offset to 10% baseline [e.g., Bandettini et al., 1993; Blamire et al., 1992; Kwong et al., 1992]. In some regions of the cortex, the response can be quite rapid. For example, Richter et al. [1996] estimated that the MR signal in the motor cortex took about 4.6 sec to reach a peak from the onset of a finger movement. This result indicates that the response in the motor cortex may be faster than that in sensory regions. Some of this difference may be due to differences in cerebral vasculature across brain regions, which can affect local delivery of oxygenated blood. We also note that the hemodynamic lag in the auditory cortex, as measured here in experiment 1, is slightly longer than that estimated by other auditory fMRI studies. For example, Hickok et al. [1997] found that the signal change in the auditory cortex peaked at approximately 5 sec poststimulus onset [see also Friston, 1997; Josephs et al., 1997]. However, the apparent difference in rise time may be explained by a difference in methodology between our study and those auditory studies mentioned above. A relatively short rise time is obtained when the hemodynamic delay is mapped using the presentation of single words (event-related fMRI) as opposed to using a short, but sustained, passage of connected speech.

As mentioned in the Introduction, the auditory response measured using event-related fMRI techniques must partly reflect an interaction between the responses to the stimulus and to the scanner acoustic noise. This interaction has not yet been explicitly characterized, but may influence the nature of the hemodynamic response in the auditory cortex. The duration of stimulation has also been shown to determine the shape of the hemodynamic response. For example, Dale and Buckner [1997] demonstrated that the hemodynamic response is approximately linear (i.e., additive across time) during the presentation of up to three single sequential events. Their data, using bursts of one, two, and three trials of a flickering checkerboard stimulus, revealed a corresponding increase in magnitude of the MR signal as the number of consecutive trials increased. This rising magnitude across time is associated with a concomitant increase in the peak lag time as the number of trials increases, due to the latency of the activation to stimuli added. The hemodynamic response cannot continue to increase in magnitude indefinitely. Data gathered using long

stimulus epochs (i.e., 30 sec or more) suggest that the MR response may plateau once a saturation point is reached (see Fig. 4) or may even decline slightly [Friston, 1997]. In addition, the estimated hemodynamic response may vary both with stimulus properties (e.g., length of stimulus and stimulus rate) and across regions of interest (e.g., different patterns of response are observed between the primary auditory cortex and Wernicke's area) [Friston, 1997; Price et al., 1992]. It may therefore not be possible to give a unique or universal value of hemodynamic delay.

Synchronization of data collection with maximum MR signal change

The primary feature of the sparse imaging technique is that one collects single sets of images near to the maximum and minimum of the hemodynamic response. It is therefore desirable to time the data acquisitions such that they optimize the difference in MR signal between stimulation and baseline conditions.

Our results from experiment 1 suggest that the absolute mean rise and fall time is around 10.5 sec. However, in experiment 2, we imaged 14 sec after the onset of each epoch, 3.5 sec *after* the maximum and minimum. The individual hemodynamic responses showed that the hemodynamic lag varied more between than within subjects. While significant individual differences might indicate adapting the procedure to optimize the MR signal change in individuals, this is not possible to justify on practical grounds. Moreover, our analysis of the hemodynamic response in the auditory cortex indicates that the change in MR signal actually *plateaus* once it has reached maximum. Therefore, a longer interscan interval, such as 14 sec, ensures that one always acquires the images near the maximum of the hemodynamic response, irrespective of the significant variations in lag time across subjects.

The mean hemodynamic response is generally asymmetrical, taking longer to return to baseline than to rise to peak. Most notably, the hemodynamic response dips below the initial baseline following stimulus offset before returning to the resting level of activation. This dip has been called ''undershoot,'' although more precisely it is a negative overshoot. fMRI data from the visual cortex suggest that the hemodynamic response may take approximately 30 sec to return to a true baseline [Hu et al., 1996]. When the auditory cortex is responding to short repeated cycles of stimulation and rest, our data from experiment 1 (Fig. 4) suggest that the experimental paradigm drives the hemodynamic response so that it becomes sinusoidal. Experiment 1 does not reveal the true baseline level of activation, as the hemodynamic response never approaches a baseline plateau during the epochs of silence. However, for the purpose of simple task subtraction, we assume an approximate equivalence between the negative overshoot and the true baseline level of activation. Indeed, by acquiring ''baseline'' data during the negative overshoot phase of hemodynamic response conditions, using sparse imaging, it is possible to optimally enhance the difference between stimulus and silence conditions.

Given the time course of hemodynamic changes in the auditory cortex, in particular the plateau of the haemodynamic peak during short periods of stimulation, some advantage may be gained by imaging at shorter intervals than the 14 sec used here. In the sparse imaging paradigm that we have described, a single volume of images is acquired at regular intervals at the transition between stimulation and baseline conditions. By reducing the duration of each epoch, it is therefore possible to increase the number of images acquired within the same total experimental time, thus increasing the number of data averages in the statistical analysis. However, since the signal-to-noise ratio of the image is maximized in other ways (e.g., by achieving equilibrium magnetization between excitations), this may afford little advantage in practice. There are several other considerations which limit the reduction in the epoch duration. If one images at the absolute point of the mean hemodynamic peak, then individual differences in the hemodynamic lag will introduce additional variation in the magnitudes of MR signal change. Thus, signal detection may be compromised across the group. Furthermore, at shorter stimulus cycles, ''baseline'' data are acquired at a time point that does not maximize the utility of the negative overshoot as a way of increasing the difference between stimulation and baseline conditions. We are currently successfully using the sparse imaging technique with an interscan interval of 11 sec in further fMRI studies of auditory activation, in order to maximize both the detection of the stimulus-induced activation and the number of data acquisitions possible in one experimental run.

ACKNOWLEDGMENTS

We thank Mark Wallace for his invaluable assistance in the anatomical classification of areas of activation.

REFERENCES

Aitkin L. 1990. The auditory cortex: structural and functional bases of auditory perception. London: Chapman and Hall.

Bandettini PA, Jesmanowicz A, Wong EC, Hyde JS. 1993. Processing strategies for time-course data sets in functional MRI of the human brain. Magn Reson Med 30:161–173.

Bandettini PA, Jesmanowicz A, Van Kylen J, Birn RM, Hyde JS. 1998. Functional MRI of brain activation induced by scanner acoustic noise. Magn Reson Med 39:410–416.

Binder JR, Rao SM, Hammeke TA, Yetkin YZ, Jesmanowicz A, Bandettini PA, Wong EC, Estowski LD, Goldstein MD, Haughton VM, Hyde JS. 1994a. Functional magnetic resonance imaging of human auditory cortex. Ann Neurol 35:662–672.

Binder JR, Rao SM, Hammeke TA, Frost JA, Bandettini PA, Hyde JS. 1994b. Effects of stimulus rate on signal response during functional magnetic resonance imaging of auditory cortex. Cogn Brain Res 2:31–38.

Blamire AM, Ogawa S, Ugurbil K, Rothman D, McCarthy G, Ellerman JM, Hyder F, Rattner Z, Shulman RG. 1992. Dynamic mapping of the human visual cortex by high-speed magnetic resonance imaging. Proc Natl Acad Sci USA 89:11069–11073.

Bowtell R, Mansfield P, Coxon RJ, Harvey PR, Glover PM. 1994. High-resolution EPI at 3.0 T. Magn Reson Mater Phys Med Biol 2:1–5.

Bullock DC, Chambers JC, Palmer AR. 1998. A high-quality sound system for use in functional magnetic resonance imaging. Br J Audiol 32:96.

Dale A, Buckner R. 1997. Selective averaging of rapidly presented individual trials using fMRI. Hum Brain Mapp 5:329–340.

Friston KJ. 1997. Imaging cognitive anatomy. Trends Cogn Sci 1:21–27.

Friston KJ, Jezzard P, Turner R. 1994. Analysis of functional MRI time-series. Hum Brain Mapp 1:153171.

Hedeen RA, Edelstein WA. 1997. Characterization and prediction of gradient acoustic noise in MR imagers. Magn Reson Med 37:7–10.

Henkelman RM, Bronskill MJ. 1987. Artifacts in magnetic resonance imaging. In: Gore JC, editor. Reviews of magnetic resonance in medicine, volume 2. New York: Pergamon Press p 1–127.

Hickok G, Love T, Swinney D, Wong EC, Buxton RB. 1997. Functional MR imaging during auditory word perception: a single-trial presentation paradigm. Brain Lang 58:197–201.

Hu X, Huu Le T, Ugurbil K. 1996. Evaluation of the early response in fMRI using short stimulus duration. Neuroimage 3:7.

Josephs O, Turner R, Friston K. 1997. Event-related fMRI. Hum Brain Mapp 5:243–248.

Kowalski N, Versnel H, Shamma SA. 1995. Comparison of responses in the anterior and primary auditory fields of the ferret cortex. J Neurophysiol 73:1513–1523.

Kwong KK, Belliveau JW, Chesler DA, Goldberg IE, Weiskoff RM, Poncelet BP, Kennedy DN, Hoppel BE, Cohen MS, Turner R, Cheng HM, Brady TJ, Rosen BR. 1992. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. Proc Natl Acad Sci USA 89:5675–5679.

Lauter JL, Herscovitch P, Formby C, Raichle ME. 1985. Tonotopic organisation in human auditory cortex revealed by positron emission tomography, Hear Res 20:199–205.

Millen SJ, Haughton VM, Yetkin Z. 1995. Functional magnetic resonance imaging of the central auditory pathway following speech and pure-tone stimuli. Laryngoscope 105:1305–1310.

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.

Price C, Wise R, Ramsay S, Friston K, Howard D, Patterson K, Frakowiak R. 1992. Regional response differences within the human auditory cortex when listening to words. Neurosci Lett 146:179–182.

Ravicz ME, Melcher JR, Talavage TM, Benson RR, Rosen BR, Kiang NYS. 1997. Characterisation and reduction of imager generated noise during fMRI. In: Association for Research in Otolaryngology, 20th Midwinter meeting, p 205.

Richter W, Ugurbil K, Kim SG. 1996. Limitations of temporal resolution in fMRI. Neuroimage 3:38.

Rinck P. 1993. Magnetic resonance in medicine. Oxford: Blackwell Scientific Publications. p 1–241.

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.

Schreiner CE, Mendelson JR. 1990. Functional topography of cat primary auditory cortex: distribution of integrated excitation. J Neurophysiol 64:1442–1459.

Shah NJ, Jäncke L, Gross-Ruyken ML, Posse S, Müller-Gärtner HW. 1997. How does acoustic masking noise affect fMRI of the auditory cortex? Neuroimage 5:195.

Shamma SA, Fleshman JW, Wiser PR, Versnel H. 1993. Organization of response areas in ferret primary auditory cortex. J Neurophysiol 69:367–383.

Sutter ML, Schreiner CE. 1991. Physiology and topography of neurons with multipeaked tuning curves in cat primary auditory cortex. J Neurophysiol 65:1207–1226

Talavage TM, Ledden PJ, Sereno MI, Rosen BR, Dale AM. 1997. Multiple phase-encoded tonotopic maps in human auditory cortex. Neuroimage 5:8.

Talavage TM, Edmister WB, Ledden PJ, Weisskoff RM. 1998a. Comparison of impact of fMRI sequence acoustics on auditory cortex activation. In: Proceedings of the 6th Annual meeting of the International Society for Magnetic Resonance in Medicine. p 1503.

Talavage TM, Edmister WB, Ledden PJ, Weisskoff RM. 1998b. Quantification of the impact of fMRI scanner noise on auditory cortex. In: Proceedings of the 6th Annual meeting of the International Society for Magnetic Resonance in Medicine. p 1502.

Wessinger CM, Buoncore MH, Kussmaul CL, Mangun GR. 1997. Tonotopy in human auditory cortex examined with fMRI. Hum Brain Mapp 5:18–25.

Woodruff PWR, Benson RR, Bandettini PA, Kwong KK, Howard RJ, Talavage T, Belliveau J, Rosen BR. 1996. Modulation of auditory and visual cortex by selective attention is modality-dependent. Neuroreport 7:1909–1913.

Zeki S. 1993. A vision of the brain. Oxford: Blackwell Scientific. p 1–366.