High Spatial Resolution Increases the Specificity of Block-Design BOLD fMRI Studies of Overt Vowel Production

David A. Soltysik[a] and **James S. Hyde**[a],[*]

a*Department of Biophysics, Medical College of Wisconsin, Milwaukee, WI 53226 USA*

Abstract

Functional MRI (fMRI) studies of tasks involving orofacial motion, such as speech, are prone to problems related to motion-induced magnetic field variations. Orofacial motion perturbs the static magnetic field, leading to signal changes that correlate with the task and corrupt activation maps with false positives or signal loss. These motion-induced signal changes represent a contraindication for the implementation of fMRI to study the neurophysiology of orofacial motion. An fMRI experiment of a structured, non-semantic vowel production task was performed using four different voxel volumes and three different slice orientations in an attempt to find a set of acquisition parameters leading to activation maps with maximum specificity. Results indicate that the use of small voxel volumes ($2\times2\times3$ mm^3) yielded a significantly higher percentage of true positive activation compared to the use of larger voxel volumes. Slice orientation did not have as great an impact as spatial resolution, although coronal slices appeared superior at high spatial resolutions. Furthermore, it was found that combining the strategy of high spatial resolution with an optimum task duration and post-processing methods for separating true and false positives greatly improved the specificity of single-subject, block-design fMRI studies of structured, overt vowel production.

Keywords

fMRI; BOLD; motion artifact; block design; speech; voxel size

INTRODUCTION

Echo planar imaging (EPI) suffers from geometric distortion, the main source of which is magnetic field inhomogeneities (Jezzard and Balaban, 1995). Static magnetic field inhomogeneities are caused by interfaces between materials of different magnetic susceptibility (such as tissue and bone or tissue and air) (Schenck, 1996). The amount of distortion increases with the strength of the magnetic field, making this a concern for high field imaging.

Despite the use of high-order shimming, large-scale inhomogeneities remain in the static magnetic field. These large-scale inhomogeneities lead to geometric distortions in the phase-encoding direction (Farzaneh et al., 1990). Researchers have sought the use of magnetic field maps to correct these geometric distortions (Jezzard and Balaban, 1995). Field maps can be

*Corresponding author. Department of Biophysics, Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, WI 53226, USA, E-mail address: jshyde@mcw.edu (J.S. Hyde), Telephone: 414-456-4005, Fax: 414-456-6512.

obtained by acquiring two separate images with different echo times (TE) and subtracting the phase images. The field map correction can be applied to any image in the time series as long as the tissue being imaged remains static. However, if tissue motion were to occur during the scan, new magnetic field inhomogeneities would be introduced, causing dynamic distortions.

It was observed by Yetkin et al. (1996) that motion of a phantom outside the field-of-view (FOV) of an image produced signal changes inside the FOV of the image. During an EPI sequence imaging a phantom, a second, smaller phantom was moved outside the head coil. Signal intensity fluctuations of the imaged phantom during static conditions were observed to be less than 1% of the baseline. However, signal intensity fluctuations between 4% and 25% of the baseline were observed during motion of the external phantom. Both positive and negative signal changes were seen in response to the small phantom motion. The greatest signal changes were seen at the edges of the phantom.

The motion of the phantom outside the FOV caused changes in the spatial configuration of the static magnetic field, B_0. These changes in B_0 extended throughout the imaging volume, creating dynamic geometric distortions. Since this effect occurred during the motion of the small phantom, signal changes in many voxels were correlated with this out-of-FOV motion. Strongly positive and negative correlations were observed between many voxel time series and the waveform representing the on/off motion pattern of the small phantom.

Functional MRI (fMRI) experiments in which subjects perform tasks involving orofacial motion, such as speech, are consequently contaminated by spurious signal changes due to dynamic magnetic field perturbations. These signal changes lead to false activation or obscure true activation during fMRI studies. An increased amount of false positives leads to low specificity, producing activation maps that are virtually useless. An early solution to this problem was to change the experiment design from block trials to event-related (ER) trials (Birn et al., 1999), in which the subject performs a short task (~1–2 s), allowing a temporal separation of motion artifact signal changes from blood oxygenation level-dependent (BOLD) response signal changes. However, event-related designs suffer from weak functional contrast-to-noise ratio (CNR) compared to block-design runs (Bandettini and Cox, 2000).

Dynamic determination of field maps has been attempted as a way to reduce geometric distortions during fMRI sequences (Birn et al., 1998; Hutton et al., 2002; Roopchansingh et al., 2003, Sutton et al., 2004). An early form of this technique required two consecutive images of the same slice to be collected with different TEs. A faster technique is to use an altered trajectory through *k*-space to acquire two images with different TEs within a single echo (Roopchansingh et al., 2003). This is repeated for all slices in the volume and all volumes in the time series, allowing the determination of a unique field map for every image. Although gross geometric distortion can be corrected via this method, movement between collections of the two images with different TEs has been shown to generate excessive temporal noise, making its use counterproductive (Hutton et al., 2002).

Birn et al. (2004) have shown that block-design and ER-design tasks with optimized durations parameters can reduce the effect of task-corrleated motion artifacts. In another recent paper (Soltysik and Hyde, 2006), it was shown that post-processing methods can be used to separate true positives caused by the BOLD response from false positives caused by motion artifacts during a block-design fMRI experiment involving the task of gum chewing. Although powerful, these technique may not remove all false positives. Furthermore, functional MRI studies of speech are likely to produce more motion artifacts than chewing, since speech requires the coordination of nearly 100 muscles (Ackermann and Riecker, 2004) as well as the controlled motion of the lungs. Greater displacement of tissue during task compared to rest means an increased amount of task-correlated signal changes. Therefore, acquisition strategies

that reduce the presence of motion-induced signal changes are of great benefit to neuroscientists interested in performing fMRI studies of overt speech production.

There is much interest in conducting fMRI studies to examine the neurological control for the articulation of speech (Lotze et al., 2000; Riecker et al., 2000; Shuster and Lemieux, 2005; Riecker et al., 2005). Thus far, these experiments have relied on event-related designs and/or subject averaging to reduce the presence of false activation caused by task-correlated orofacial motion. These strategies are not ideal, however, and they fail to produce artifact-free activation maps from individual subjects. The goal of the present study was to determine whether there exists an optimum voxel size and slice orientation that reduce the presence of false positives caused by magnetic field perturbations. These optimum acquisition strategies, combined with an optimum task duration and post-processing methods, could then be used to obtain activation maps with high specificity for block-design fMRI experiments of single subjects performing tasks involving orofacial motion.

There has been much research on the selection of voxel size in fMRI (Frahm et al., 1993; Thompson et al., 1994; Hoogenraad et al., 1999; Howseman et al., 1999; Yoo et al., 2001; Hyde et al., 2001). As voxel volume is decreased, there are two conflicting effects (Howseman et al., 1999): decreased image SNR (Macovski, 1996) and increased functional CNR (Frahm et al., 1993; Thompson et al., 1994; Hoogenraad et al., 1999). The former is a consequence of less signal being available from smaller voxels. The latter results from the minimization of partial volume averaging of gray matter, where BOLD signal changes occur, and other tissue like white matter and cerebrospinal fluid, where BOLD effects do not occur. Yoo et al. (2001) found that BOLD CNR is maximized when the spatial resolution matches the size of the activation. Hyde et al. (2001) concluded that a voxel volume of $1.5 \times 1.5 \times 1.5$ mm^3 was optimum to address the issues of gray matter tortuosity and partial volume averaging. Therefore, increasing spatial resolution, up to a certain point, is expected to improve the detection accuracy of fMRI. However, the effect of spatial resolution on the specificity of activation maps from fMRI studies involving orofacial motion has not been previously examined.

The optimization of slice orientation in fMRI studies has not received a lot of attention in the literature, although one study has suggested that oblique axial slices yield the best signal detection accuracy for a motor cortex experiment and coronal slices yield the worst (Gustard et al., 2001). The spatial dimensions of the human brain in the Talairach atlas (Talairach and Tournoux, 1988) (omitting the cerebellum) are 105 mm (I/S), 165 mm (A/P), and 123 mm (R/L). Thus, axial slices yield the greatest volume of brain coverage for a given temporal resolution. However, tasks like speech may induce head motion along the superior-inferior plane or about the right-left axis, i.e., pitch. To improve volume registration during runs involving head motion, it is prudent to acquire slices such that the expected motion would occur in-plane and not through-plane, giving volume registration routines the chance for adequate correction. Through-plane motion results in a loss of spin-excitation histories. If motion is expected to occur only in the superior-inferior plane, then coronal or sagittal slices would be best. If motion is expected to occur in the superior-inferior plane and along the right-left axis, then sagittal slices would be best. It is unknown if the presence of the orofacial muscles in the image FOV would increase the presence of motion artifact signal changes in the brain. Considering all of these factors, it is not clear which slice orientation would be preferred for fMRI studies of speech.

MATERIALS AND METHODS

Subjects

Fifteen healthy, right-handed, native-English speakers (9 males, 6 females, aged 32 ± 10 years) were recruited from the local community. The subjects were screened for contraindications such as metal implants or claustrophobia. All subjects reviewed and signed a written consent form approved by the Human Research Review Committee at the Medical College of Wisconsin.

Experiment Design

Each subject performed structured, overt vowel production tasks during twelve block-design runs. The structured speech consisted of a periodic alternation between two vowel sounds. Each run began with 6 s of rest followed by ten epochs. Each epoch consisted of 12 s of task followed by 12 s of rest. The 12 s duration for both task and rest is considered optimum as this creates a 90° temporal phase difference between motion artifact signal changes and the BOLD response (Soltysik and Hyde, 2006). During the task period, the subject was presented with the visual cue of "a" to phonate the long vowel sound /e/ as in "ate" or the visual cue of "e" to phonate the long vowel sound /i/ as in "see." Each phonation cue was displayed for 857 ms, allowing for 14 phonation periods, alternating between "a" and "e," in the 12 s task block.

In order to articulate the /i/ sound, the tongue is high in the mouth and fronted, nearly filling the oral cavity. The artifactual image distortion that occurs when this vowel is sounded is expected to be particularly severe. We are concerned with speech-correlated tissue displacements that modulate the magnetic field in the brain because of the magnetic susceptibility of air and tissue. The tongue is somewhat lowered to articulate the /e/ sound, and the lips and jaw may move. Alternating production of vowel sounds as instructed by the visual cue is expected to require attention by the subject, and the result is stronger fMRI response than repeated production of the same sound. The task-design subtracts the brain response during alternating sound production of the two vowels from the brain response with articulatory muscles held in a neutral position during rest. This design was expected to present an extreme technical challenge: true fMRI response and artifactual response from out-of-field-of-view correlated motion of tissues were expected to be nearly the same in a block trial paradigm. The goal of the study was to develop acquisition strategies that would permit separation of the resulting true and false positives. The task is deliberately intended to be a "worst case scenario."

In addition, this task was expected to activate the motor cortex responsible for articulation but not to activate cortices involved with language. Although this task does not mimic the rapid transitions or the large changes in the oral cavity often present in normal speech, it is still expected to produce severe motion artifacts in the image data. Subjects were instructed to phonate the speech sounds overtly while avoiding rigid head motion. Subjects were also instructed to remain silent and still during the rest periods, which were visually cued by a cross-hair. The visual presentation was programmed in E-Prime (Psychology Software Tools, www.pstnet.com/) and projected onto a translucent screen at the foot of the MRI table. A prism on the head coil allowed the subject to view the screen. Each run lasted 4 minutes and 6 seconds, leading to a 123 image volumes collected per run. Foam pads were used to minimize head motion.

Data Acquisition

Data were acquired on a 3T GE Long Bore Excite MRI scanner (General Electric Healthcare, Waukesha, WI), using an 8-channel array RF head coil. After the scout scan, a T_1-weighted spoiled GRASS (SPGR) sequence (TI/TE/FA = 450 ms/3.2 ms/12°, FOV = 24 cm, number of slices = 124, slice thickness = 1.2 mm) was performed to obtain high-resolution images of the

whole brain. Next, a high-order autoshimming sequence was performed. Following this, EPI anatomic volumes were acquired covering the whole brain (TR/TE/FA = 4000 ms/30 ms/87°). The fourth and final shots of these short runs were used as base volumes during volume registration of the functional EPI data sets. These EPI anatomic volumes were acquired at four different spatial resolutions (FOV = 25.6 cm, matrix 128×128, slice thickness = 3 mm, number of slices = 43; FOV = 19.2 cm, matrix 64×64, slice thickness = 4 mm, number of slices = 35; FOV = 25.6 cm, matrix 64×64, slice thickness = 4 mm, number of slices = 35; and FOV = 25.6 cm, matrix 64×64, slice thickness = 5 mm, number of slices = 28). This protocol resulted in acquisitions with four different voxel dimensions ($2 \times 2 \times 3$ mm^3, $3 \times 3 \times 4$ mm^3, $4 \times 4 \times 4$ mm^3, and $4 \times 4 \times 5$ mm^3).

The vowel production tasks were performed during EPI sequences (TR/TE/FA = 2000 ms/30 ms/77°) that were acquired in three different slice orientations (axial, coronal, and sagittal) and four different voxel dimensions ($2 \times 2 \times 3$ mm^3, $3 \times 3 \times 4$ mm^3, $4 \times 4 \times 4$ mm^3, and $4 \times 4 \times 5$ mm^3) representing four different voxel volumes (12 µl, 36 µl, 64 µl, and 80 µl). For the coronal and sagittal orientations, the frequency-encoding direction was chosen to be in the superior/inferior direction to reduce the effect of signal dropout near the sinuses. Thus, twelve functional runs were acquired. The order of the functional runs was determined pseudorandomly for each subject. Partial brain volumes were acquired in each case. For each run, 20 slices were acquired. The axial acquisition began approximately 1 cm from the top of the brain, extending inferiorly, and covering the left and right primary sensorimotor cortices and part of the supplementary motor area (SMA). The coronal acquisition was positioned midway along the anterior/posterior axis of the brain, also covering the left and right primary sensorimotor cortices and the SMA. The sagittal slices were acquired in two lateral groups of ten, covering the left and right primary sensorimotor cortices.

Data Analysis

Data were analyzed using AFNI (Cox, 1996) and locally written programs in MATLAB (The Mathworks, Natick, MA). The first three image volumes (or 6 s of images) of the functional EPI data sets were discarded to allow the equilibration of magnetization. For each EPI data set, the fourth shot of the whole brain EPI anatomic volume with the matching voxel size was used as the base volume during volume registration. Since the EPI anatomic volumes were acquired with axial slices, they had to be resampled to match the specific voxel dimensions of the coronal and sagittal EPI data sets. The functional EPI data sets were also zeropadded to match the spatial extent of the EPI anatomic volumes. This volume registration technique was performed to correct displacement of the brain along the slice select axis.

The high-resolution whole brain image set was warped to Talairach space (Talairach and Tournoux, 1988), after which a skull-stripping algorithm was implemented. The expected ROI mask was drawn on the brain of one subject in Talairach space by designating three regions of interest (ROI) where activation for speech articulation was expected (Fig. 1). The first ROI expected to activate during articulation was the middle portion of the primary sensorimotor cortex (Paus et al., 1996; Indefrey and Levelt, 2000; Lotze et al., 2000; Riecker et al., 2000; Shuster and Lemieux, 2005; Riecker et al., 2005). This ROI included the right and left pre- and postcentral gyri, extending from z = 20 mm to 60 mm. The second ROI expected to activate was the SMA (Paus et al., 1996; Indefrey and Levelt, 2000; Lotze et al., 2000; Riecker et al., 2000; Riecker et al., 2005). This ROI was drawn from the brain vertex, superiorly, to the cingulate sulcus, inferiorly, from y = 1 mm to y = −24 mm, and from x = −16 mm to x = 16 mm (Chainay et al., 2004). The third ROI expected to activate was the left insula (Wise et al., 1999; Riecker et al., 2000; Shuster and Lemieux, 2005; Ackermann and Riecker, 2004; Riecker et al., 2005). This ROI was drawn around the gray matter region bounded by the temporal lobe and frontal lobe, extending from z = −6 mm to z = 21 mm. The identification of cortical regions

was assisted by the use of the Talairach Daemon database in AFNI. The expected ROI mask created for one subject was used for the other subjects, after brains were warped to Talairach space.

A multiple regression analysis was performed on all voxels in the functional EPI data sets, fitting the baseline to a second order polynomial and using a BOLD response waveform as the reference function. This waveform was formed by convolving the task reference function with a gamma-variate impulse function. Activation was thresholded at $F > 21.35$ ($P < 1 \times 10^{-5}$).

Three motion-suppression thresholds were then applied to the activation maps (Soltysik and Hyde, 2006). These included the correlation-phase threshold, the noise-to-baseline ratio (NtBR) threshold, and the cluster threshold.

For this experiment, a slightly-modified version of the correlation-phase threshold was used in which voxels were accepted for which the temporal phases were within 45° of the BOLD response waveform. The phase of the BOLD response waveform, $r[t]$, was determined by first calculating the fast Fourier transform,

$$R[f] = \text{FFT} \{r[t]\},$$

(1)

and using a 4-quadrant arctangent function to compute the phase,

$$\varphi = \arctan_4(-\text{Im}\{R[f_T]\}, \text{Re}\{R[f_T]\}),$$

(2)

where we used the real and imaginary parts of the Fourier transform of the reference function evaluated at the task frequency, f_T, to compute the phase of the reference function, φ_{ref}.

The temporal phase of each statistically significant voxel was then determined by evaluating (Eq. (1)) and (Eq. (2)) using that voxel time series in place of $r[t]$. Activated voxels with phases within the range $[\varphi_{ref} - 45°, \varphi_{ref} + 45°]$ were accepted, and voxels with phases outside of this range were rejected. This threshold rejected voxels with a temporal phase indicative of a motion artifact response.

The NtBR threshold accepted only those voxels for which the NtBR value was less than 1, where we define:

$$\text{NtBR} = \frac{\sigma_{\text{res}}}{S_{\text{baseline}}},$$

(3)

$$\sigma_{\text{res}} = \left[\frac{1}{N(K-1)} \sum_{n=1}^{N} \sum_{k=1}^{K} \left(\overline{S}_n - S_{nk} \right)^2 \right]^{1/2},$$

(4)

where the signal had a value S at time point n and epoch k for a total of N time points and K epochs. This threshold essentially rejected voxels that yielded a trial-averaged response with a large amount of variance compared to the baseline.

The cluster threshold accepted only those voxels found in a cluster of activated voxels, the minimum size of which was determined by Monte Carlo calculations to yield an overall statistical threshold of $\alpha = 0.01$. This threshold removed remaining false positives that were isolated after the previous thresholds. It also handled the issue of multiple comparisons.

The activation remaining after the statistical threshold and the three motion-suppression thresholds was designated the *total activation volume*, or V_{TOT}. This volume contained both true and false positives accepted by the thresholds.

Next, the EPI data sets containing activation maps were warped to Talairach space and resampled to 1-mm-cubic voxels. In this new coordinate system, the total activation volume was multiplied by the expected ROI mask. The activation located inside the expected ROIs remained, while the activation located outside the expected ROIs was eliminated. This resulted in what was designated the *true positive activation volume*, or V_{TP}. The volume eliminated is termed the *false positive activation volume*, or V_{FP}. The *percent of activation that was true positive*, or P_{TP}, was then calculated by dividing the true positive activation volume by the total activation volume,

$$P_{TP} = \frac{V_{TP}}{V_{TOT}} \times 100\% = \frac{V_{TP}}{V_{TP} + V_{FP}} \times 100\%.$$

(5)

Equation (5) implies that as the false positive activation volume decreases, the percent of activation that is true positive will increase, making P_{TP} proportional to specificity, which is defined as:

$$\text{specificity} = 1 - \text{FPF} = 1 - \frac{V_{FP}}{V_F},$$

(6)

where V_F represents the total false volume (including false positives and true negatives). It should be noted that specificity also increases as the false positive activation volume decreases.

RESULTS

Block-design fMRI experiments of overt speech production are notorious for yielding poor activation results. Rigid head motion during speech can reduce functional CNR of the BOLD response due to spatial displacement of active tissue. Movement of the speech articulators can create dynamic field perturbations, leading to geometric distortion and increased temporal noise. Changes in the size and susceptibility of the air cavity in the lungs during exhalation can produce magnetic field inhomogeneities, leading to signal loss (Raj et al., 2001). Differences among the subjects' speech performances, which were not recorded, could lead to differences in activation results. Thus, the quality of activation was first examined by averaging P_{TP} across all runs for each subject. Values of P_{TP} were only included in this calculation if V_{TOT} was nonzero. Despite poor expectations, it was found that nine subjects performed well, with average values of P_{TP} ranging from 31% to 54%, while six subjects performed poorly, with average values of P_{TP} ranging from 0% to 13%. The six poorly activating subjects were excluded from further analysis.

For each case of voxel volume and slice orientation, the values of P_{TP} were averaged across the remaining nine subjects. Values were excluded, however, if V_{TOT} or V_{TP} were zero. A zero value for either variable was considered indicative of a bad run.

Table 1 shows values of P_{TP} for each good subject and each set of parameters. Plots of P_{TP} versus voxel volume are shown in Fig. 2 for the axial, coronal, and sagittal orientations. Linear regressions were performed to find the line that best fit the data. These data show that progressively higher percentages of true positive activation volume result from using smaller voxel volumes. If the true positive activation volume is expected to remain constant, this implies that the false positive activation volume decreases with voxel volume. This is shown to be true for all three slice orientations.

To illustrate the effect more clearly, a two sample *t* test was performed to see if there was a significant difference in P_{TP} between the acquisition for the smallest voxel size (12 µl) and the acquisition for the largest voxel size (80 µl) (see Fig. 3). The 12 µl voxel case had a significantly higher P_{TP} than the 80 µl voxel case for the axial ($P = 0.02$), coronal ($P = 0.00002$), and sagittal ($P = 0.007$) acquisitions.

Next, the value of P_{TP} was compared across slice orientations for the 12 µl voxel size. The coronal acquisition had a significantly higher P_{TP} than the axial acquisition ($P = 0.004$). The coronal acquisition had a somewhat significantly higher P_{TP} than the sagittal acquisition ($P = 0.06$). No significant difference was seen between P_{TP} values for axial and sagittal acquisitions ($P > 0.1$). Furthermore, when comparing slice orientations for the 80 µl voxel size, no significant differences were seen between any two P_{TP} values ($P > 0.1$).

Activation maps for a single subject are shown on 3D rendered brain images in Fig. 4. It can be seen that the three runs using small voxel volume acquisitions (Fig. 4, top row) all show narrow strips of activation, mostly confined to the motor strip located along the precentral gyrus. The three runs using large voxel volume acquisitions (Fig. 4, bottom row) show clusters of activation extending past the motor strip as well as false positive activation outside the brain. A greater consistency is also found for the activation maps achieved using small voxels compared to those found using large voxels. The activation volumes appear larger for the large voxel acquisitions, but the specificity, or the ability to reject the null hypothesis, is visibly worse when compared to the results for the small voxel acquisitions.

DISCUSSION

The data shown in Fig. 2, Fig. 3, and Fig. 4 reveal that high spatial resolution reduces the percentage of activation that is false positive. Furthermore, it was shown that the combined use of high spatial resolution, an optimum task duration, and post-processing methods could result in activation maps with a high specificity for block-design fMRI experiments on single subjects performing structured, overt vowel production tasks.

The six subjects in this study that were excluded from further analysis produced very poor or no activation in the expected areas of activation. Rigid head motion for these subjects was not found to be excessive when compared with the other subjects. Therefore, these subjects may have employed speaking strategies that produced strong magnetic field perturbations. These field perturbations could have led to signal loss and greatly reduced functional CNR. Speech involves not only motion of orofacial tissue, such as tongue, lips, and jaw, but also a controlled movement of the lungs and diaphragm. Bulk susceptibility variations in the lungs during respiration has been shown to produce signal loss in brain tissue being imaged (Raj et al., 2001). Differences in inhalation strategies may have led to different results among the subjects. Training subjects to speak by exhaling small volumes of air may reduce the risk of respiration-induced signal loss during fMRI studies of overt speech. Further research needs to be done to understand the problem of false negatives in fMRI studies of overt speech.

Figure 5 suggests that increased amounts of true positive activation volume result from using larger voxel volumes. However, the definition of true positive volume used in this analysis included all of the activation surviving the thresholds and located inside the expected ROIs. Volume averaging of active and non-active tissue would have exaggerated the true positive activation volume for the larger voxel acquisitions. Thus, we cannot conclude from Fig. 5 that large voxel acquisitions will necessarily give us more true positive activation volume.

The data do not suggest an optimum slice orientation across all voxel volumes. Figure 3 indicates that coronal slices are better than axial or sagittal slices when 12 µl voxels are used, but no slice orientation appears to be superior when 80 µl voxels are used. Since overall head motion during the functional runs was generally observed to be less than 2 mm, it is reasonable to conclude that an optimum slice orientation is necessary only for small voxel acquisitions. For the acquisitions using small voxel volumes, the data revealed that the coronal acquisition yielded a significantly higher P_{TP} than both axial and sagittal acquisitions. However, if speech

is likely to produce head motion mostly in the superior/inferior plane, it is unclear why the sagittal acquisition would not perform just as well as the coronal acquisition.

The structured, overt vowel production task used in this experiment does not reflect the rapid, dynamic nature of normal speech. Movement of more articulators such as the lips and tongue will induce increased amounts of spurious signal changes in images. Faster overt vowel production may cause more extreme motion artifactual signal changes. Rapid speech can be expected to result in high frequency artifactual noise that could, in some paradigms, be more easily filtered from true positives characterized by the rather sluggish fMRI hemodynamic response. The structured, overt vowel production task of this paper seems likely to produce more false-positive activation than would a true speech production task.

Chewing tasks involve lateral-medial motion, which is different than the sagittal plane motion of speech. Future studies are needed to determine if the strategies of high spatial resolution and coronal acquisition can improve specificity for chewing tasks as well as overt speech tasks that use more articulators.

There are several sources of error in this experiment. The identification of true positive activation for each subject is likely to contain errors. It was postulated that accepting only thresholded activation lying in ROIs expected to activate during articulation (sensorimotor cortex, SMA, and left insula) would be an adequate means of identifying true positive activation. However, any false positive activation remaining in these ROIs would be falsely labeled true positive. Conversely, any true positive activation not occurring inside these ROIs would be ignored. The expected ROIs were purposely drawn large to avoid the latter error, but that increases the likelihood of the former error. Increased amounts of false positives mislabeled as true positives would have led to higher values than expected for P_{TP}.

Another source of error is that the motion-suppression thresholds may have removed some true positive activation. This could have led to lower values than expected for P_{TP} if the true positives removed were inside the expected ROIs or higher values than expected for P_{TP} if the true positives removed were outside the expected ROIs. The inability to define clearly the region of true positive activation makes inferences about sensitivity difficult. However, maximizing specificity was the main goal of this experiment, so a good estimate of true positive activation was sufficient for our purposes.

Image warping due to static field inhomogeneities is a problem for echo-planar imaging, and can vary across slice orientation and spatial resolution (Haacke et al., 1999). This could have affected the results of the experiment. Haacke et al. (1999) describe how increasing the read gradient (and hence the spatial resolution) can reduce geometric distortions cause by local magnetic field inhomogeneities. It is possible that there is a relation between the reduced geometric distortion and the improvement of specificity seen for small voxel acquisitions.

Although the results of this paper reveal that the use of small voxels and coronal slices is best for fMRI studies of structured, overt vowel production, there are disadvantages associated with the use of these parameters. The main disadvantage is that they require a sacrifice in either temporal resolution or spatial coverage of the brain. High resolution images require longer imaging times than lower resolution images when larger matrix sizes are used. In addition, more coronal slices are required to cover the whole brain compared with axial slices. To meet the optimum acquisition parameters in this study and cover the whole brain, an EPI acquisition would require 56 3-mm-thick coronal slices with a matrix size of 128×128 and FOV of 256 mm. Assuming each slice can be acquired in 90 ms, a TR of about 5 s would be required. This temporal resolution may not be suitable for some experiments. Furthermore, increasing the gradient strength of the imaging sequence results in an increased acoustical output from the EPI pulse sequence, which may be uncomfortable to some subjects. Alternatively, the

researcher may decide to use axial slices without fearing too much loss of specificity. The whole brain (minus the cerebellum) could be covered with 43 3-mm-thick axial slices using a TR of 4 s. If higher temporal resolution is required, then acquisition of a partial brain volume would be recommended.

CONCLUSIONS

A block-design fMRI experiment was run in which subjects performed a structured, overt vowel production task. EPI runs were acquired with four different voxel volumes and three different slice orientations in an attempt to find the optimum acquisition parameters that maximize specificity. Results indicated that the use of small voxel volumes ($2\times2\times3$ mm^3) yielded a significantly higher percentage of true positive activation compared to the use of larger voxel volumes. The disadvantage of high spatial resolution is that a sacrifice must be made in either temporal resolution or brain coverage. In addition, spatial resolution was found to be a more important factor to specificity than the slice orientation. However, at high spatial resolution, coronal slices were found to work best. Combining the strategy of high spatial resolution with an optimum task duration and post-processing methods to separate true and false positives has been found to greatly improve the specificity of single-subject, block-design fMRI studies of structured, overt vowel production.

ACKNOWLEDGMENTS

REFERENCES

Ackermann H, Riecker A. The contribution of the insula to motor aspects of speech production: A review and a hypothesis. Brain and Language 2004;89:320–328. [PubMed: 15068914]

Bandettini PA, Cox RW. Event-related fMRI contrast when using constant interstimulus interval: theory and experiment. Magn. Reson. Med 2000;43:540–548. [PubMed: 10748429]

Birn RM, Bandettini PA, Cox RW, Jesmanowicz A, Shaker R. Magnetic field changes in the human brain due to swallowing or speaking. Magn. Reson. Med 1998;40:55–60. [PubMed: 9660553]

Birn RM, Bandettini PA, Cox RW, Shaker R. Event-related fMRI of tasks involving brief motion. Hum. Brain. Mapp 1999;7:106–114. [PubMed: 9950068]

Birn RM, Cox RW, Bandettini P. Experimental designs and processing strategies for fMRI studies involving overt verbal responses. NeuroImage 2004;23:1046–1058. [PubMed: 15528105]

Chainay H, Krainiki A, Tanguy M-L, Gerardin E, Le Bihan D, Lehericy S. Foot, face, and hand representation in the human supplementary motor area. NeuroReport 2004;15:765–769. [PubMed: 15073511]

Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput. Biomed. Res 1996;29:162–173. [PubMed: 8812068]

Farzaneh F, Riederer S, Pelc NJ. Analysis of T2 limitations and off-resonance effects on spatial resolution and artifacts in echo-planar imaging. Magn. Reson. Med 1990;14:123–139. [PubMed: 2352469]

Frahm J, Merboldt K-D, Hanicke W. Functional MRI of human brain activation at high spatial resolution. Magn. Reson. Med 1993;29:139–144. [PubMed: 8419736]

Gustard S, Fadili J, Williams EJ, Hall LD, Carpenter TA, Brett M, Bullmore ET. Effect of slice orientation on reproducibility of fMRI motor activation at 3 Tesla. Magn. Reson. Imag 2001;19:1323–1331.

Haacke, EM.; Brown, RW.; Thompson, MR.; Venkatesan, R. Magnetic Resonance Imaging: Physical Principles and Sequence Design. New York: John Wiley & Sons, Inc.; 1999. p. 569-617.

Hoogenraad FGC, Hofman MBM, Pouwels PJW, Reichenbach JR, Rombouts SARB, Haacke EM. Sub-millimeter fMRI at 1.5 Tesla: correlation of high resolution with low resolution measurements. J. of Magn. Reson. Imaging 1999;9:475–482. [PubMed: 10194720]

Howseman AM, Grootoonk S, Porter DA, Ramdeen J, Holmes AP, Turner R. The effect of slice order and thickness on fMRI activation data using multislice echo-planar imaging. NeuroImage 1999;9:363–376. [PubMed: 10191165]

Hutton C, Bork A, Josephs O, Deichmann R, Ashburner J, Turner R. Image distortion correction in fMRI: A quantitative evaluation. NeuroImage 2002;16:217–240. [PubMed: 11969330]

Hyde JS, Biswal BB, Jesmanowicz A. High-resolution fMRI using multislice partial k-space GR-EPI with cubic voxels. Magn. Reson. Med 2001;46:114–125. [PubMed: 11443717]

Indefrey, P.; Levelt, WJM. The neural correlates of language production. In: Gazzaniga, MS., editor. The New Cognitive Neurosciences. Cambridge: MIT Press; 2000. p. 845-865.

Jezzard P, Balaban RS. Correction for geometric distortion in echo planar images from B_0 field variations. Magn. Reson. Med 1995;34:65–73. [PubMed: 7674900]

Lotze M, Seggewies G, Erb M, Grodd W, Birbaumer N. The representation of articulation in the primary sensorimotor cortex. NeuroReport 2000;11:2985–2989. [PubMed: 11006980]

Macovski A. Noise in MRI. Magn. Reson. Med 1996;36:494–497. [PubMed: 8875425]

Paus T, Perry DW, Zatorre RJ, Worsley KJ, Evans AC. Modulation of cerebral blood flow in the human auditory cortex during speech: role of motor-to-sensory discharges. Eur. J. Neurosci 1996;8:2236–2246. [PubMed: 8950088]

Raj D, Anderson AW, Gore JC. Respiratory effects in human functional magnetic resonance imaging due to bulk susceptibility changes. Phys. Med. Biol 2001;46:3331–3340. [PubMed: 11768509]

Riecker A, Ackermann H, Wildgruber D, Dogil G, Grodd W. Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. NeuroReport 2000;11:1997–2000. [PubMed: 10884059]

Riecker A, Mathiak K, Wildgruber D, Erb M, Hertrich I, Grodd W, Ackermann H. fMRI reveals two distinct cerebral networks subserving speech motor control. Neurology 2005;64:700–706. [PubMed: 15728295]

Roopchansingh V, Cox RW, Jesmanowicz A, Ward BD, Hyde JS. Single-shot magnetic field mapping embedded in echo-planar time-course imaging. Magn. Reson. Med 2003;50:839–843. [PubMed: 14523971]

Schenck JD. The role of magnetic susceptibility in magnetic resonance imaging: MRI magnetic compatibility of the first and second kinds. Med. Phys 1996;23:815–850. [PubMed: 8798169]

Shuster LI, Lemieux SK. An investigation of covertly and overtly produced mono- and multisyllabic words. Brain and Language 2005;93:20–31. [PubMed: 15766765]

Soltysik DA, Hyde JS. Strategies for block-design fMRI experiments during task-related motion of structures of the oral cavity. NeuroImage 2006;29:1260–1271. [PubMed: 16275020]

Sutton BP, Noll DC, Fessler JA. Dynamic field map estimation using a spiral-in/spiral-out acquisition. Magn. Reson. Med 2004;51:1194–1204. [PubMed: 15170840]

Talairach, J.; Tournoux, P. Co-planar Stereotaxic Atlas of the Human Brain. New York: Thieme Medical; 1988.

Thompson RM, Jack CR, Butts K, Hanson DP, Riederer SJ, Ehman RL, Hynes RW, Hangiandreou NJ. Imaging of cerebral activation at 1.5 T: optimizing a technique for conventional hardware. Radiology 1994;190:873–877. [PubMed: 8115643]

Wise RJS, Greene J, Buchel C, Scott SK. Brain regions involved in articulation. Lancet 1999;353:1057–1061. [PubMed: 10199354]

Yetkin FZ, Haughton VM, Cox RW, Hyde J, Birn RM, Wong EC, Prost R. Effect of motion outside the field of view on functional MR. Am. J. Neuroradiol 1996;17:1005–1009. [PubMed: 8791907]

Yoo S-S, Guttmann CRG, Panych LP. Multiresolution data acquisition and detection in functional MRI. NeuroImage 2001;14:1476–1485. [PubMed: 11707104]
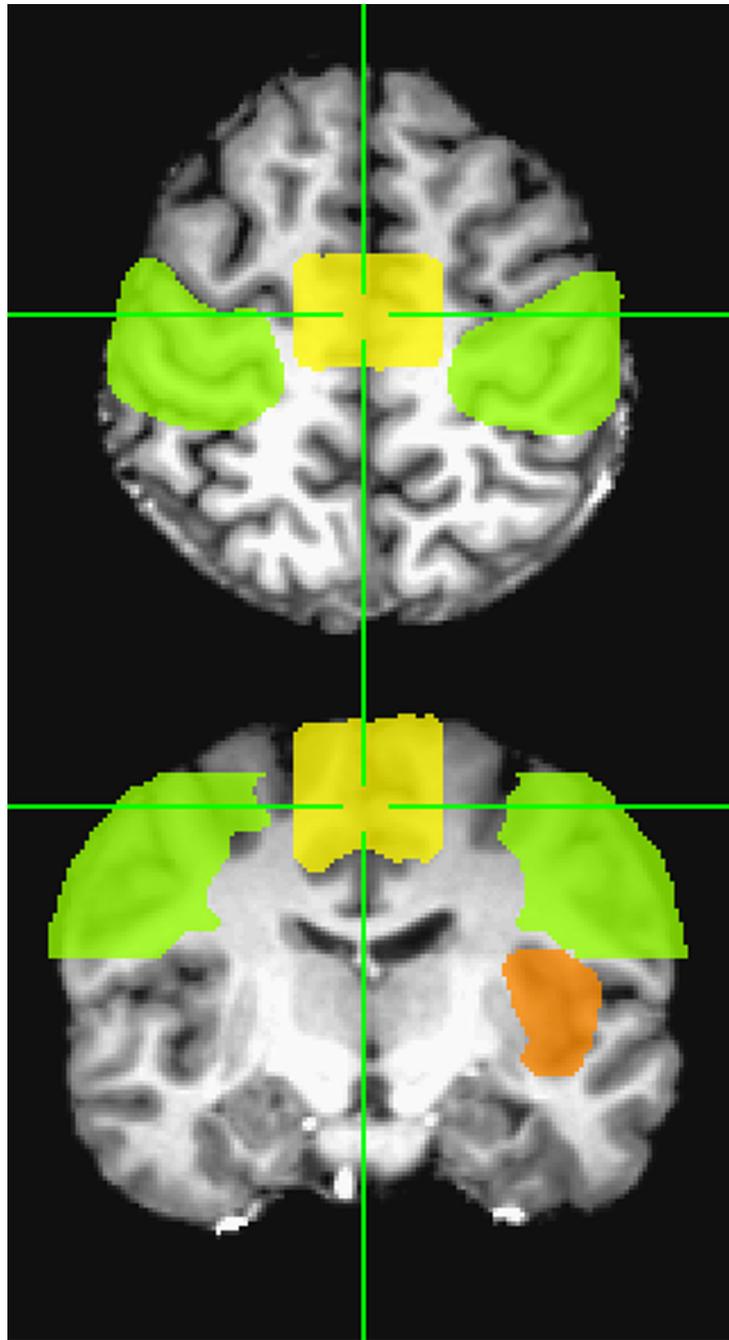
Fig. 1.
The expected ROI consisted of the primary sensorimotor cortex (green), the supplementary motor area (yellow), and the left insula (orange) shown here in axial and coronal cross sections.

P$_{TP}$ (Axial Acquisition)

P$_{TP}$ (Coronal Acquisition)

P$_{TP}$ (Sagittal Acquisition)
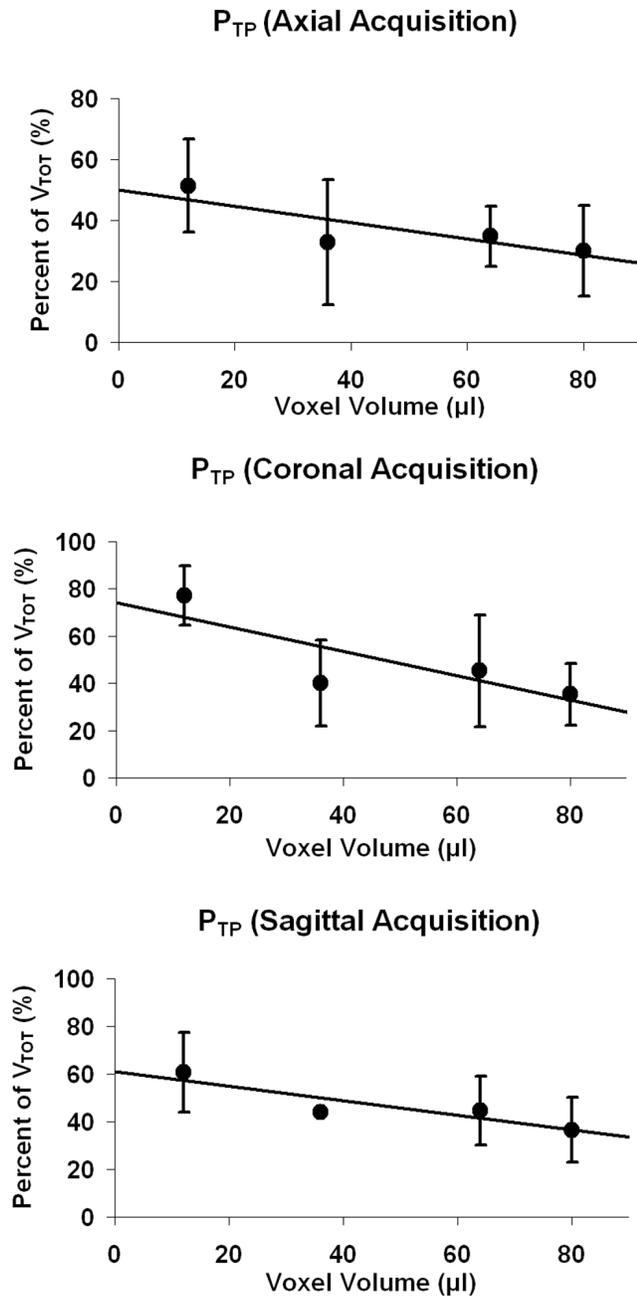
Fig. 2.
Plots of percent of activation volume that is true positive, P_{TP}, across voxel volume for axial, coronal, and sagittal acquisitions. A decrease in P_{TP} is seen for increasing voxel volume in all three cases. (Error bars represent standard deviation.)
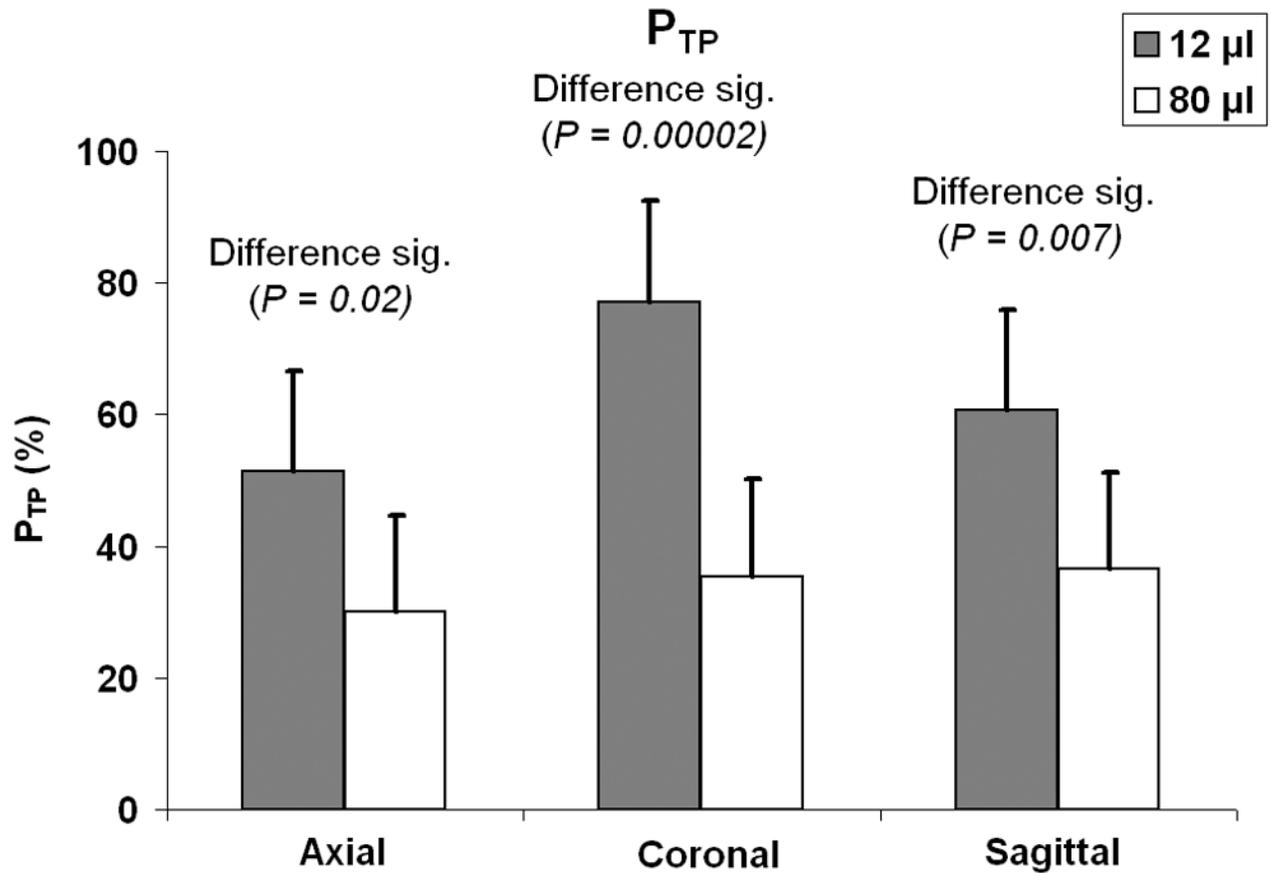
Fig. 3.
Comparison of the percent of activation volume that is true positive, P_{TP}, for the acquisition with the smallest voxel volume (12 µl) with the acquisition with the largest voxel volume (80 µl) for all three slice orientations. A two-sample t test was used to determine the significance of the difference resulting from voxel volume. (Error bars represent standard deviation.)
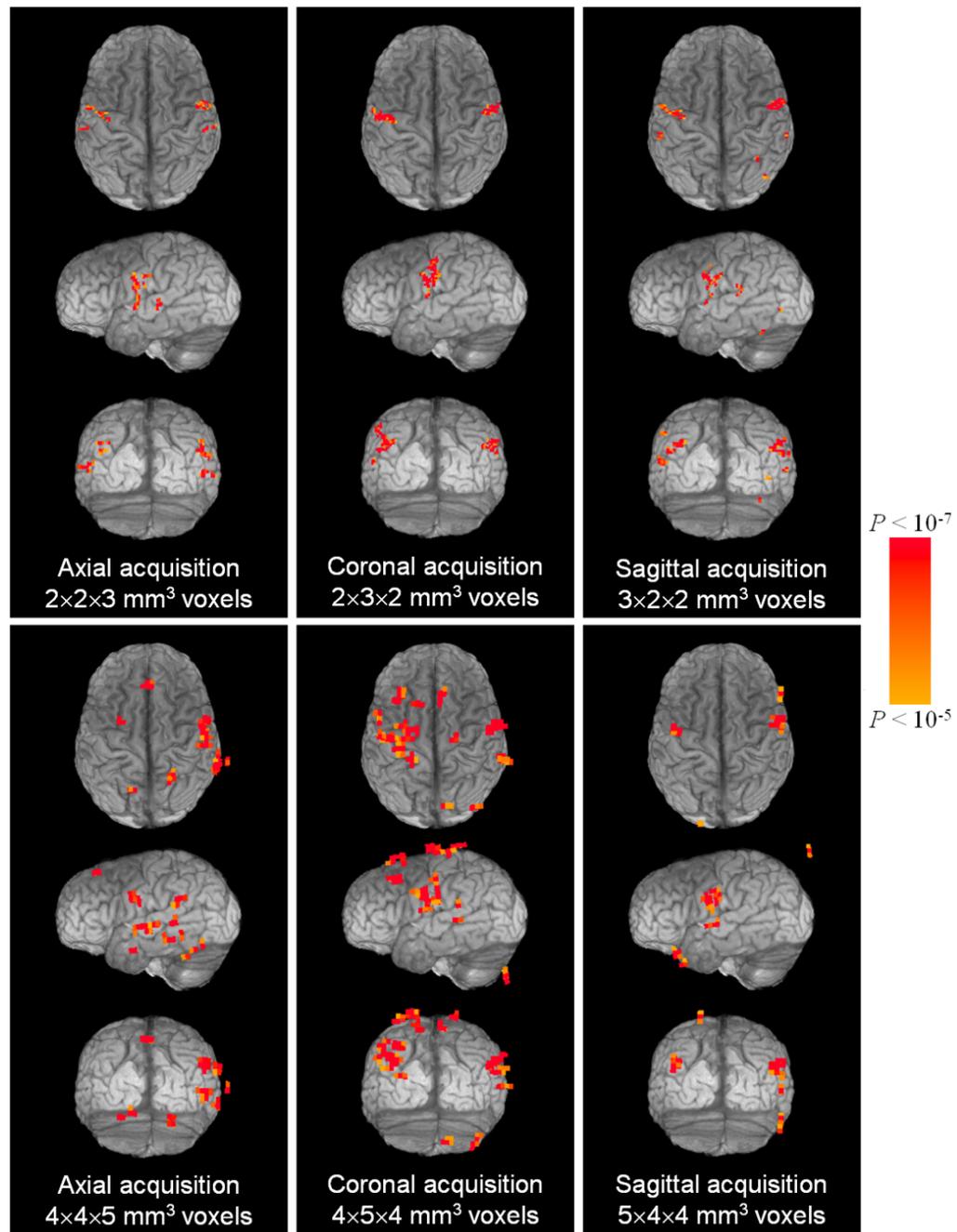
Fig. 4.
For a single subject, 3D rendered images of the brain are shown transparent to activation (V_{TOT}) remaining after statistical and motion-suppression thresholds ($\alpha = 0.01$). (Voxel dimensions are given in terms of R-L, A-P, and S-I axes, respectively. On axial and coronal views, right is on the right.)
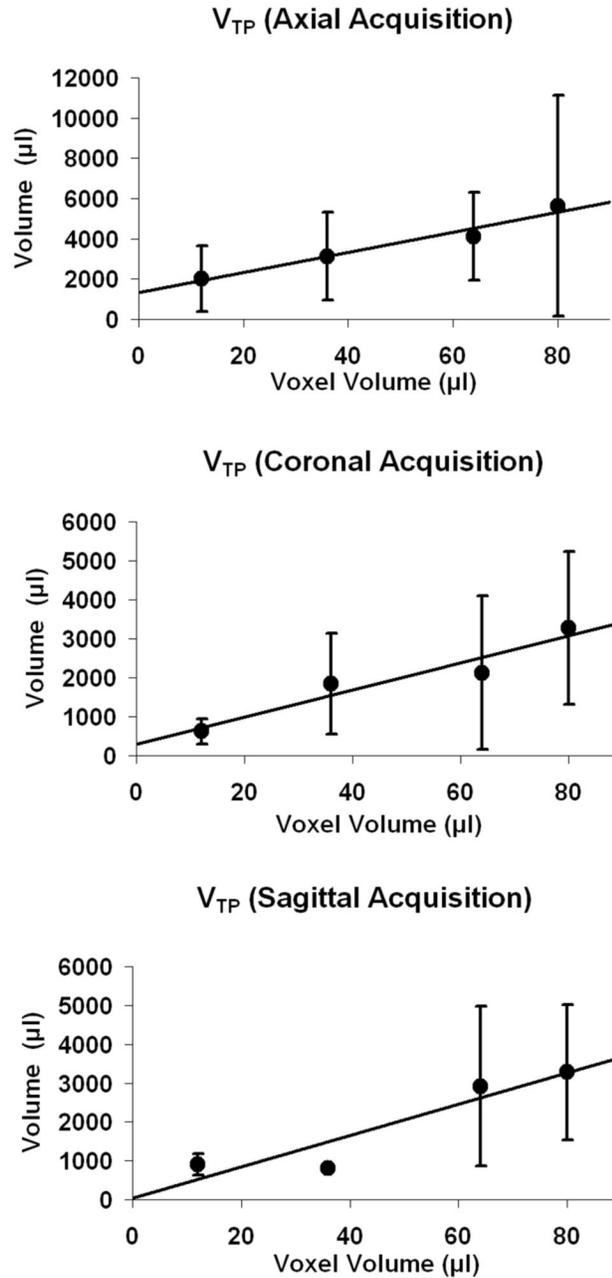
Fig. 5.
Plots of activation volume that is true positive, V_{TP}, across voxel volume for axial, coronal, and sagittal acquisitions. An increase in V_{TP} is seen for increasing voxel volume in all three cases. (Error bars represent standard deviation.)

Table 1

Percent of activation volume that is true positive, P_{TP}, (%)

AXIAL

Voxel Volume (µl)	sub02	sub05	sub06	sub10	sub11	sub12	sub13	sub14	sub15
12	66.42	43.70	65.09	-	55.86	53.15	25.68	35.10	66.05
36	11.57	34.22	37.71	18.29	75.12	49.57	19.04	37.17	12.8
64	20.51	41.85	42.39	35.71	41.43	25.79	49.70	26.45	28.88
80	17.69	24.57	27.56	64.82	15.25	33.20	37.40	23.51	26.00

CORONAL

Voxel Volume (µl)	sub02	sub05	sub06	sub10	sub11	sub12	sub13	sub14	sub15
12	94.04	67.54	71.11	-	94.09	-	64.44	80.87	67.93
36	20.19	32.03	59.29	-	57.42	51.36	42.85	48.18	8.82
64	41.69	18.16	54.41	-	92.09	30.45	62.01	34.89	28.26
80	26.89	29.50	45.15	46.26	36.94	37.97	53.21	34.13	8.42

SAGITTAL

Voxel Volume (µl)	sub02	sub05	sub06	sub10	sub11	sub12	sub13	sub14	sub15
12	-	43.00	72.98	-	66.48	61.42	31.98	76.19	72.42
36	-	-	-	44.00	-	-	-	-	-
64	54.01	35.71	11.23	50.46	61.62	49.19	41.28	48.54	48.32
80	23.26	40.01	39.27	33.99	54.51	17.77	22.61	40.95	56.01