

The Anatomy of Auditory Word Processing: Individual Variability

Martha W. Burton,* Douglas C. Noll,† and Steven L. Small*

**Department of Neurology, University of Maryland School of Medicine; and †Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI*

This study used functional magnetic resonance imaging (fMRI) to investigate the neural substrate underlying the processing of single words, comparing activation patterns across subjects and within individuals. In a word repetition task, subjects repeated single words aloud with instructions not to move their jaws. In a control condition involving reverse speech, subjects heard a digitally reversed speech token and said aloud the word “crime.” The averaged fMRI results showed activation in the left posterior temporal and inferior frontal regions and in the supplementary motor area, similar to previous PET studies. However, the individual subject data revealed variability in the location of the temporal and frontal activation. Although these results support previous imaging studies, demonstrating an averaged localization of auditory word processing in the posterior superior temporal gyrus (STG), they are more consistent with traditional neuropsychological data, which suggest both a typical posterior STG localization and substantial individual variability. By using careful head restraint and movement analysis and correction methods, the present study further demonstrates the feasibility of using overt articulation in fMRI experiments. © 2001 Academic Press

Key Words: neuroimaging; neuroanatomy; language; lexicon; brain mapping; functional magnetic resonance imaging (fMRI); auditory perception; articulation.

INTRODUCTION

Functional neuroimaging with positron emission tomography (PET) and functional MRI (fMRI) has recently shown promise for answering a number of outstanding questions about the neurological localization of language. Although the method of lesion analysis in patients with cerebral injury, particularly stroke, has led to much understanding of neurobiological systems involved in language processing, these newer neuroimaging approaches have increasingly addressed more specific questions, at a finer resolution, and with fewer confounding influences (e.g., from impaired patients and their recovery) than has been possible with lesion analysis methods alone.

The support of the National Institutes of Health under Grant NIH DC R01-3378 is gratefully acknowledged. We are grateful to Karalyn Patterson for providing word lists. We further appreciate the help of a number of people who contributed to this project and manuscript: Darren Emge, Robert Cox, Petr Hlustik, Craig Mullins, Ana Solodkin, and Greg Zaorski. Data from this article were presented at the 5th annual meeting of the Cognitive Neuroscience Society, San Francisco, California, in April, 1998. Steven L. Small is now at the University of Chicago Department of Neurology.

Address correspondence and reprint requests to Martha W. Burton, Department of neurology, University of Maryland, Baltimore, 12-011 Bressler Research Building, 655 W. Baltimore Street, Baltimore, MD 21201-1559. Fax: (410) 706-0324. E-mail: burton@cortex.neurology.umaryland.edu.

Auditory single-word processing is one such area of language processing that is the focus of this report. During auditory presentation of words compared to rest, regional cerebral blood flow (rCBF) analyses of normal subjects using positron emission tomography (PET) have generally shown activity in the bilateral posterior superior and middle temporal gyri (Démonet et al., 1992; Petersen et al., 1988, 1989; Price et al., 1996; Wise et al., 1991). Much of this activation has been attributed to processing of the auditory signal because similar patterns of activation have been shown for nonwords and nonspeech stimuli (O'Leary et al., 1996; Wise et al., 1991). Nevertheless, when speech and nonspeech tasks are directly compared, several studies have shown left superior temporal activation that may be due to semantic processing and/or increased attentional demands in the processing of complex speech stimuli (Grady et al., 1997; Howard et al., 1992; O'Leary et al., 1996; Pugh et al., 1996; Small et al., 1999).

The localization of auditory word processing to the posterior portion of the left superior temporal gyrus (STG) reflects an average pattern of activation among many subjects (Howard et al., 1992). This pattern is consistent with classic neuropsychology and the notion that language comprehension involves the posterior part of the STG. Although it is possible that the subjects in previous studies were homogeneous in their rCBF responses to the experimental task, it is even more likely that they were heterogeneous. Such heterogeneity has been typical in lesion analysis studies of lexical processing, leading to extensive philosophical debate in the field of aphasia research regarding the validity of group versus single-subject research studies (Caramazza, 1986). Recent fMRI and PET studies of individual subjects on a wide variety of tasks have reinforced the heterogeneity argument, showing that the functional anatomy of individual brains differs significantly, particularly in associational cortex (e.g., language) (Steinmetz & Seitz, 1991; Van Essen et al., 1998). An important goal of the present study was to understand the nature of this variability, i.e., the individual contributions to this "grand average" pattern of activation. Thus, using a pair of tasks developed for one of these PET studies (Howard et al., 1992), the purpose of this study was to investigate individual variability in the functional neuroanatomy of auditory word processing.

Despite the overall consistency of word processing studies, some variation in the location of activation during auditory word processing tasks has been noted (e.g., Fiez et al., 1996; Howard et al., 1992; Price et al., 1996). Activity in temporoparietal regions may depend on whether the words that are presented are phonologically distinct (e.g., telephone and airplane), whether the subject is asked to perform a discrimination task based on temporal speech cues, and whether the words are repeated during the scan (Fiez et al., 1996). Activation of the operculum of the inferior frontal gyrus, which also occurs in some word repetition tasks, may also depend on specific task requirements (including decision components that are thought to be eliminated by the "subtraction" method), rate of presentation, and image analysis technique (Price et al., 1996); thus, its role(s) in language perception and production remain enigmatic (Gabrieli, Poldrack, & Desmond, 1998).

Some of the difficulty ascertaining structure/function relationships for particular areas of the brain comes from an incomplete understanding of the cognitive processes involved in particular language tasks. When compared to rest conditions, complex tasks often activate large networks, making it difficult to associate specific brain regions with particular cognitive processes (Small et al., 1999). Furthermore, when control conditions include passive listening or rest, there is potentially a great deal of variability in a subject's level of attention and mental activity during that phase of the experiment.

To control this attentional aspect of subject performance, some studies have used carefully constructed condition pairs in which an active condition and a control condition differ by only a single cognitive “component” according to some information processing models of the system (Newell & Simon, 1972). One such experiment aimed to dissect out the receptive component of lexical phonology by contrasting two active word processing tasks (Howard et al., 1992). In the word repetition task, subjects repeated single words aloud. In the reverse speech task, subjects heard a digitally reversed speech token and said the word “crime.” This word was selected because it has characteristics (i.e., frequency, imageability, number of phonemes, and number of syllables) that are at the average of the words used in the repetition task.

Both tasks were subjected to careful analysis of cognitive subcomponents (Howard et al., 1992). The repetition task was thought to consist of (a) primary auditory processes, (b) complex pattern recognition, (c) auditory input lexicon, (d) semantic access, (e) word retrieval, (f) word production, (g) and articulatory programming. The reverse speech task was said to incorporate the same input and output processes with the exception of semantic access and word retrieval components. Thus, by “subtracting” one task from the other (Donders, 1868/1969), the results of the experiment were thought to demonstrate the neuroanatomical sites of auditory word processing. Typically, such processing has been associated with the posterior superior temporal gyrus, an area that has been implicated in lexical processing for over a century (Wernicke, 1874).

It is possible to criticize this particular task comparison for a number of reasons. Although it is necessary to attend to the auditory input to perform successfully the articulation step in the word condition, this is not necessary in the reverse speech condition. However, it is becoming clear that reverse speech, which has the same amount of acoustic variability as the speech signal, is more intelligible as speech than had been previously realized (Saberri & Perrott, 1999). Another difference is that the articulated word is novel in the word condition but not in the control condition. Thus, the comparison would control adequately articulation, but perhaps not the lexical phonological processing aspects of the speech. This is only a problem under an assumption that input and output phonological processes are distinct, which is not necessary to the present study.

Several reservations about the method of task comparison, or “subtraction” (Burton & Small, 1999; Friston et al., 1996; Sergent et al., 1992; Sternberg, 1969), have been cited, including the difficulty involved in identifying the cognitive processes involved in particular tasks and the generally incorrect assumptions about interactivity and feedback relationships among subprocesses. Nevertheless, the method provides a direct way to compare the present data with previously acquired data that have strong historical foundations.

Although overt articulation with fMRI is more complicated than with PET due to the necessity of coregistering over 100 individual images, fMRI can provide high anatomical precision in individual subjects. By contrast, most previous studies of lexical access and storage involved analysis of group PET data with spatial smoothing and intersubject averaging (Howard et al., 1992; Price et al., 1996). By using fMRI, the present study can explore the individual variability in the anatomy of auditory word processing and shed light on the variability in localization that is found through neuropsychological study and increasingly through disparate neuroimaging results (e.g., see Poeppel, 1996; Demonet et al., 1996, for discussion).

Tasks which require overt articulation in the MRI scanner provide a particular challenge for fMRI because of known serious problems with head motion artifacts (Hajnal et al., 1995). One solution to this problem has been to instruct subjects to

respond silently during word repetition and word generation tasks (Rueckert et al., 1994; Wildgruber et al., 1996). In a direct comparison of overt and covert speech tasks, Yetkin et al. (1995) showed that generating words aloud or silently both produce activation, though in their study significantly more activation and fewer artifacts were found in the silent condition (see also Barch et al., 1999). Previously, we demonstrated that reliable activation patterns can be obtained when subjects speak aloud in the scanner (Small et al., 1996), but those subjects used a "bite bar" to restrict severely head movement. Bite bars are often uncomfortable for subjects, but without such head restraint, movement artifact may increase. Another recent approach has been to use single-trial design to reduce artifacts by distinguishing average responses from a brief stimulus from signal response due to motion (Birn et al., 1999). Nonetheless, most fMRI studies still do not use overt articulation because of unresolved problems with motion artifacts. In the present study, subjects were trained to keep their jaws immobilized to minimize head movement and all other muscles of articulation were free to move.

METHODS

Subjects

Fourteen monolingual native speakers of American English (10 female and 4 male, ranging in age from 22 to 49, mean = 30) participated in the study. Thirteen of the subjects were strongly right handed, with scores greater than +50 on the Edinburgh Handedness Inventory (Oldfield, 1971). The other subject's score was -7.7. (We had not restricted our hypotheses to right-handed subjects and so did not exclude this subject). The study was approved by the University of Maryland School of Medicine Institutional Review Board and all subjects provided written informed consent prior to participation.

Stimuli

Subjects performed the auditory word processing tasks described by Howard et al. (1992). In the word repetition task, subjects heard a sequence of spoken words and were asked to repeat each of the words aloud without moving their jaw. Although no muscles of articulation were actually restrained physically and all were free to move, the subjects were taught how to speak without jaw movement and practiced for a few minutes prior to scanning. In the "hear and say" condition, which served as a control, subjects heard a series of words that had been digitally reversed. These reversed word stimuli were created from the actual word stimuli to control for any possible duration effects. Subjects spoke the word "crime" in response to each reverse word, also without moving the jaw muscles. "Crime" was selected for the response word in the "hear and say" task by Howard et al. (1992) because its characteristics (length = 4 phonemes, frequency = 34, and imagery = 471) were similar to the average of those in the word lists.

The word repetition stimuli consisted of 540 words selected from the MRC psycholinguistic database (Coltheart, 1981). The words were one or two syllables in length with the number of phonemes ranging from two to eight. The average frequency of the word lists was 82 words per million (Francis & Kucera, 1982) and average imageability was 491. These words were divided into three lists matched for frequency, number of syllables, number of phonemes, and rated imageability. These word lists were derived by adding to the original lists, without changing their average characteristics.

Procedures

Subjects were placed in the scanner and head movement was restricted with specially contoured foam rubber pillows. Air-conduction headphones (Resonance Technologies, Northridge, CA) were placed around the ears and connected to a stereo system controlled by a Macintosh computer. The Macintosh computer ran the PsyScope psychological software system (Cohen et al., 1993) to present the experimental conditions. Subjects were made as comfortable as possible with the headphones and pillows prior to scanning, an effort resulting in decreased head motion. Stimulus presentation was organized in sequential blocks, beginning with word repetition followed by "hear and say." Each block lasted 24 s. The word repetition task consisted of a tone to indicate the beginning of a task followed by a word presented

every 1500 ms in random order. Presentation of reverse speech was identical except that instead of a word, subjects heard the reverse speech tokens. The sequence of word repetition followed by "hear and say" was repeated 12 times, for a total of 576 s (9 min 36 s) per experimental trial. This entire experimental trial was performed three times, each time with a different list of words.

Image acquisition. The functional imaging studies were carried out in the Department of Radiology at the University of Maryland Medical Center on a 1.5-T Signa scanner (GE Medical Systems, Milwaukee, WI) with a standard GE quadrature head coil. Twenty-four 6-mm structural T1-weighted anatomical images [500 ms repetition time (TR), 9 ms gradient echo time (TE), spin echo pulse sequence] were acquired in the sagittal plane to determine the anatomy of the functional slices, starting from the left temporal lobe and ending with the right. Functional data were acquired using the spiral k-space method (Noll et al., 1995). A single shot gradient echo spiral scan pulse sequence provided 3.24 mm in plane resolution interpolated to 1.88 mm over a 24-cm field of view (FOV). T2*-weighted imaging was accomplished with a TE of 35 ms and a TR of 4000 ms with a flip angle of 60°. A complete set of the 24 slice locations was generated every repetition time cycle (4 s). Each of the 24 slices was acquired 6 times during each 24-s task interval. Since each task was repeated 12 times during each experimental trial, the total number of images obtained for each condition for each experimental trial was 72. Following the end of each trial, a brief (<1-min) break occurred, during which the next trial was prepared. Each trial was repeated 3 times to improve the statistical power and reliability, yielding three times series of data, each containing 72 time points per voxel location. Thus, for each subject, there were 216 images for word repetition and 216 images for "hear and say."

Image analysis. Prior to statistical analysis of the functional data, several preliminary image processing steps were applied to the data following reconstruction of the spiral k-space (Fourier space) images. To compensate for signal drift, each time series was normalized to a zero mean and a linear detrending procedure was applied to remove changes in the baseline signal. To correct for head movement by the subject, all the images for each subject were coregistered using the 3D linear (six-parameter) automated image registration algorithm (Woods, Cherry, & Mazziotta, 1992). The three resulting time series (one from each experimental trial) were concatenated together. Due to acquisition error, the data from one subject (subject 1) contained a single experimental trial and another (subject 2) two trials. Thus, these two time series were shorter than the other 12.

Cross-correlation thresholding was then performed on the concatenated time series (Bandettini et al., 1993) using AFNI software (Cox, 1996). In this analysis, the concatenated waveform for each voxel was correlated with a reference waveform corresponding to the temporal pattern of stimulus presentation. The cross-correlation analysis was performed with sine waves offset by 0, 7.5, 15, 22.5, and 30 degrees, corresponding to time delays of 0, 2.5, 5, 7.5, and 10 s (to account for hemodynamic lag), and the best match retained.

A separate oropharyngeal motion analysis was performed on the correlated data because of an *a priori* decision to exclude subjects with excessive oropharyngeal movement during overt articulation. Four strongly right-handed subjects (two female and two male) showed substantially more activity than the other subjects in the oropharyngeal region [greater number of clusters of active voxels (see below) located below the inferior aspect of the cerebellum than the mean + 2 SD of the other 10 subjects (Mean = 7.7, SD = 7.8)]. Thus, these four subjects were excluded from further data analysis, and subsequent phases of analysis were performed on the remaining 10 subjects.

To compare these data with those obtained in previous PET studies of lexical processing, in particular the studies of Howard et al. (1992) and Price et al. (1996), the individual data were subjected to a group analysis according to the methods commonly in use for PET studies. This involved transforming the images into a standard coordinate system, smoothing them, and then averaging across subjects.

To this end, the images were first transformed into the most commonly used stereotactic coordinate space, using a series of linear transformations (Talairach & Tournoux, 1988). The images were then smoothed using a full-width half-maximum Gaussian filter 16 mm wide (cf. Price et al., 1996) and averaged after applying a Fisher transformation to the correlation coefficients. The Fisher transformation normalizes the data to compensate for between-subject differences in the correlation values. The theoretical basis for this averaging is that the arctanh of the sample correlation coefficient is approximately normally distributed, unlike the correlation coefficients themselves. Peak areas of activation were located by using a 3D clustering program with a connectivity radius of 1.9 mm, a volume threshold of 42 mm³, or 2 voxels. The functional data were plotted on an average of the structural images from the 10 subjects after transformation into Talairach coordinates.

In addition to comparing group data from fMRI and PET, and assessing the feasibility of performing fMRI studies involving overt articulation, it was important to understand the individual subject contributions to the overall group result. To achieve this aim, the group data were further analyzed as individual cases, with careful note of the variability in pattern and location of activation. The correlation coefficient data for each subject were thresholded at 0.35 and cluster sizes were measured after applying a minimum threshold (using a 3D neighborhood of 26 voxels) of greater than 2 voxels (volume = 42 mm³). For

each of the clusters, the Euclidean distance in millimeters from the center of the averaged data in that region was calculated to provide a measure of variability of the location of brain activation (e.g., the distance from each cluster in the temporal lobe to the mean Talairach coordinates of the centers of mass of the three clusters in the temporal lobe from the averaged data). The distances were then averaged and the standard deviation was calculated.

RESULTS

Group Data

Figure 1 shows the comparison of activation during the word repetition task and the "hear and say" task averaged across 10 subjects. The average data indicated three areas of peak change in activation located on the infero-lateral surface of the posterior portion of the left superior temporal gyrus (see top panel of Fig. 1). The centers of mass were located at stereotactic coordinates: ($x = -65$ mm, $y = -38$ mm, $z = 7$ mm), ($x = -63$ mm, $y = -46$ mm, $z = 12$ mm), and ($x = -57$ mm, $y = -54$ mm, $z = 15$ mm). These clusters are consistent with previous imaging studies showing posterior temporal activation (Howard et al., 1992; Price et al., 1996) and with the classic neuropsychology of auditory word processing (Dejerine, 1891). In addition, the supplementary motor cortex (SMA) near the border of pre-SMA and SMA proper produced a large cluster of activation (see bottom panel of Fig. 1). Peak activation for this cluster was located at ($x = -3$ mm, $y = 3$ mm, $z = 60$ mm), about 3–4 mm different in each plane from the location reported by (Price et al., 1996). A fifth smaller cluster that narrowly missed the cluster size threshold was located on the superior aspect of left inferior frontal gyrus near the inferior frontal sulcus separating the inferior and middle frontal gyri ($x = -52$ mm, $y = 13$ mm, $z = 27$ mm), about 4 mm more lateral and 15 mm superior to the inferior frontal gyrus location reported in Experiment 2 of Price et al. (1996) (see middle panel of Fig. 1). No other clusters approached threshold.

Individual Subject Analyses

The location of clusters of peak temporal and frontal activation for each subject are plotted in a schematic diagram of a left lateral parasagittal slice in Fig. 2, along with the peak activation reported in the (Price et al., 1996) study, which includes a reanalysis of the Howard et al. data using more sensitive PET analysis methods and a replication of the Howard et al. study with four subjects. (Talairach coordinates and volumes for the clusters in each area for each of the individual subjects are shown in Table 1). The general anatomic location of the pattern of response was similar among subjects. Eight of the 10 subjects showed strong changes in activation along the superior temporal gyrus (STG). Of those subjects, the location of the peak activation of only one (subject 2) was substantially different from that of the group average data, varying from the fMRI center of activation by 48 mm primarily because of its anterior location. Of the two subjects without STG activation, one (subject 6) showed only SMA activation and the other (subject 7) did not have any clusters that reached the .35 threshold. The precise location of clusters indicating STG activation differed among the subjects (mean distance of clusters from center of average activation = 17 mm, standard deviation = 12 mm), although a strong region of activation nonetheless appeared in the group data.

In addition to the STG activation seen in both the PET and fMRI studies, four of the subjects (4, 8, 9, and 10) revealed significant activation changes in the inferior frontal gyrus. As in the temporal lobe, the distance between the location of clusters of activity in individual subjects in the frontal lobe and the group average data were

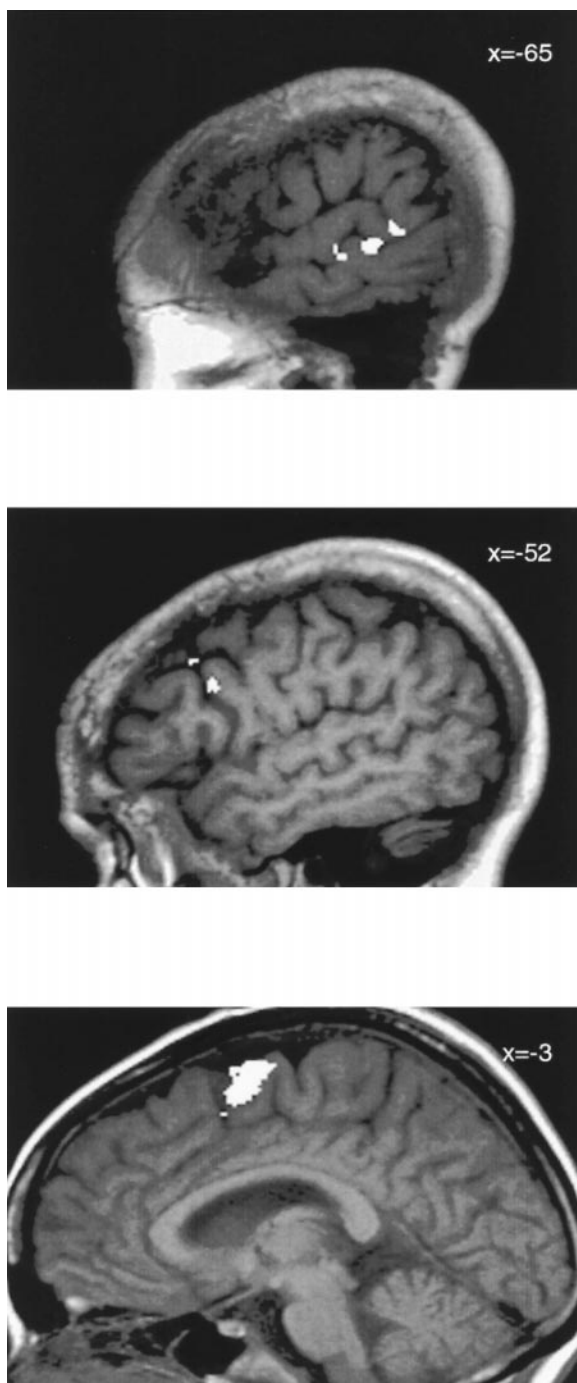


FIG. 1. Peak activation in left superior temporal gyrus (top panel), left inferior frontal gyrus (middle panel), and supplementary motor area (bottom panel) in group averaged data across subjects shown in sagittal view. The numbers in the top left corner of each panel (in white) indicate the stereotaxic x coordinates. Data are plotted on structural images of an individual subject after transformation into Talairach and Tournoux (1988) coordinates.

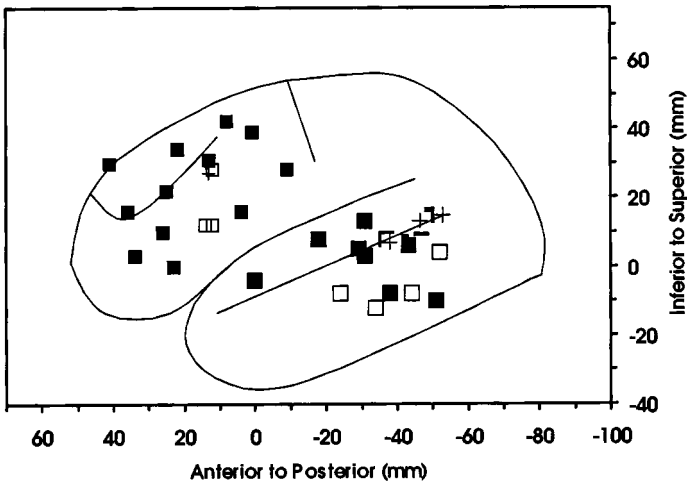


FIG. 2. Plot of location of peak temporal and frontal activity from the Price et al. (1996) PET study (open squares) and activity in individual subjects from present study (solid squares). The “+” indicates center of mass for group averaged data. Data (y and z stereotactic coordinates) are plotted on a schematic diagram of a left parasagittal slice. The horizontal axis is the y coordinate in Talairach space, indicating anterior/posterior location, and the vertical axis is the z coordinate in Talairach space, indicating inferior/superior location. The line in the temporal lobe represents the superior temporal sulcus and the curved line in the frontal lobe represents the inferior frontal sulcus.

calculated. Although the amount of variability in the frontal lobe appears to be greater than that of the temporal lobe, the mean distance from center of average activation (mean = 20 mm, standard deviation = 8 mm) was similar to that of the temporal lobe. An unpaired t test comparing distance of clusters in the temporal lobe to distance in the frontal lobe showed that this difference was not significant [$t(23) = .73, p > .47$]. An additional t test excluding the anterior temporal activation of Subject 2, which was beyond 2 standard deviations from the mean, did not reach significance [$t(22) = 1.6, p > .11$].

Five subjects (3, 4, 6, 8, and 9) showed activation in the supplementary motor area (SMA). The amount of individual variability in SMA (mean distance from center of average activation = 9 mm, standard deviation = 5 mm) was approximately half that of the temporal and frontal lobes. In addition, three subjects (1, 4, and 7) displayed right cerebellar activation. Thus, not only did the subjects vary in the location of their functional activation within brain regions, other areas in some of the subjects were active during the word task compared to the reverse speech.

To determine whether the inclusion of one left-handed subject (subject 4) affected the pattern of results, her data were compared to the other subjects. The location of activation for subject 4 was similar to the other subjects with strong left superior temporal, inferior frontal, and supplementary motor activation. However, this subject also activated additional areas with several clusters in the right cerebellum and a small cluster in the right superior temporal gyrus. Only one other subject (9), whose laterality index was +50, showed any right STG activation. Thus, although the location of activation in the three main regions (left STG, left IFG, and SMA) was similar, the left-handed subject activated the most regions and included right activation, which was only seen in one other subject.

DISCUSSION

The results of the present study lead to three fundamental conclusions. First, with anticipated loss of some data due to oropharyngeal motion, it is possible to perform

TABLE 1
Talairach Coordinates (x, y, and z) and Cluster Size of Activation in Each Region for Each Subject

	Temporal				Lateral frontal				Mesial frontal			
	x	y	z	vol	x	y	z	vol	x	y	z	vol
1	-51	-29	5	506								
1	-64	-31	13	84								
2	-57	0	-4	63								
3	-58	-51	15	801	-65	-9	28	190 ^a	-1	19	51	169
3					-51	1	39	105 ^a				
4	-60	-37	8	654	-53	26	10	506	-4	0	63	338
4	-59	-18	8	400	-46	36	16	485	-4	-7	69	127
4	-45	-38	-8	126 ^b	-46	41	30	147				
4					-53	25	22	127				
4					-53	22	34	105				
5	-60	-50	15	42								
6									-3	2	57	127
7												
8	-53	-31	3	253	-53	13	31	1181	-7	-1	60	443
8	-59	-46	12	232	-53	23	0	253	-7	-3	68	127
8	-60	-51	-10	84 ^b	-53	34	3	84				
9	-67	-43	6	84	-55	8	42	170	-12	-1	62	42
10	-67	-41	7	84	-61	4	16	42				

Note. The correlation threshold is .35. Volume (vol) of cluster sizes is given in cubic millimeters.

^a Activation is located in precentral gyrus of the frontal lobe. All other activation is in the inferior frontal gyrus, middle frontal gyrus, or on the border between inferior and middle frontal gyrus.

^b Activation is located in inferior temporal lobe. All other activation is in superior temporal gyrus or border between superior and middle temporal gyrus.

studies involving overt articulation using functional MRI. Motion artifacts resulting from even 1 mm of movement are well known to occur in fMRI studies even in the absence of overt subject movement (Frackowiak et al., 1997). However, in the present study, only 28% of subjects were excluded on the basis of movement, and they were only excluded on the basis of an objective scoring method for oropharyngeal motion. The second conclusion is that the average location of increased BOLD response in fMRI on lexical access corresponds well to that of increased rCBF with PET and to the expected left posterior superior temporal gyrus implicated in neuropsychological lesion analysis. The third main conclusion is that the data contributing to this grand average pattern of activation are heterogeneous and that there is substantial individual variability in brain activation on this task. Given the similarity in the overall (average) pattern between PET and fMRI, this variability in BOLD response is potentially true of the rCBF response as measured by PET as well.

The grand average results of the present study replicate the previous PET studies that originated the task comparison used here (Howard et al., 1992; Price et al., 1996). Eight of the subjects and the group data indicate peaks of activation in the left posterior superior and middle temporal lobe, with the peak activation in the group average occurring in the posterior STG. PET activation on this task was found at the junction of the superior and middle temporal gyri immediately inferior to the primary auditory

cortex (Howard et al., 1992). A replication and reanalysis of these data reconfirmed this approximate location, with peak activation in posterior middle temporal gyrus (Price et al., 1996). The areas identified in both the present study and the PET studies are all close to or fall within areas characteristically implicated in lexical processing by a variety of methods over a long history.

Despite the fact that the global average data from the present fMRI study do mirror the PET data (Howard et al., 1992; Price et al., 1996), the main foci of temporal activation differ by about 2 cm according to the two methods. The fMRI activation was both lateral (about 10 mm) and superior (about 15 mm) to that of the PET activation. For technical reasons, recent data suggest that fMRI activation may be slightly displaced in the direction of venous drainage from a cortical region (Kinahan & Noll, 1999). Although this is complicated in the temporal lobe, with three large veins leaving the posterior lateral temporal lobe, in the superior (Trolard), postero-inferior (Labbe), and anterior (Middle Cerebral) directions, they all route along the cortical surface. Thus, some lateral displacement of activation might be predicted. The more superior location of the temporal activation could be a function of venous drainage or of the particular subjects chosen.

Alternatively, the activation could be more reflective of activity in the classic Wernicke's area. Although a precise anatomical definition of Wernicke's area is controversial, more restrictive definitions have confined it to the posterior third of the superior temporal gyrus (Mesulam, 1990). The activation of most of the subjects in the present study falls closer to that area than the PET centers of activation in the middle and inferior temporal regions (Price et al., 1996). However, it should be noted that inferior temporal activation in the fMRI study could be reduced as a result of a susceptibility artifact (an area where the MR signal is absent) just anterior to the expected region of activation.

The other cortical and cerebellar regions of activation in the PET studies were present in at least one of the fMRI subjects, with variable contributions to the group average. The two most important such regions in the fMRI data are the operculum of the inferior frontal gyrus and the supplementary motor area. Four of the subjects showed activation in the inferior frontal gyrus (IFG). Interestingly, such activation was not present in the original PET study (Howard et al., 1992), but appeared in a reanalysis of that study using more sensitive image analysis techniques (Price et al., 1996). Overall, IFG activation is less robust than that of the STG, with a smaller area activated in the average data. Word repetition compared to reverse speech may activate this frontal area less consistently than other tasks such as verbal fluency, which involve several other cognitive processes such as selection of semantic knowledge from working memory that also are believed to activate left prefrontal cortex (Gabrieli, Poldrack, & Desmond, 1998). The present study does not clarify the circumstances under which the IFG is involved in lexical processes (cf. Dronkers, 1998 for a discussion of the role of Broca's area in language processes). However, it does suggest that in addition to task differences resulting in inconsistent activation of Broca's area, even with the same task, there are individual differences and that the amount of variability in the frontal regions is similar to that of the temporal regions.

Five subjects produced activation in the supplementary motor area. This area also produced a cluster in the analysis of group data. These results are consistent with other studies of auditory word processing (Chee et al., 1999; Petersen et al., 1988; Price et al., 1996). SMA activation seems more likely to be found in tasks which involve speech production. It might be expected that because subjects are producing words in both the repeat and reverse tasks (either repeating what they heard or saying "crime") motor activation should not emerge. However, the difference between the novelty of repeating a series of different words and repeatedly producing the same

word may be responsible for significant differences in activation between the two tasks (cf. Price et al., 1996). Interestingly, the variability in location of center of mass of SMA clusters was about half that found in the frontal and temporal lobes. The greater variability in the temporal and frontal lobes could be due to greater variability in the functional anatomy in association cortices or could be due to the smaller number of clusters found in SMA.

The PET studies observed changes in several other areas which did not appear in the averaged fMRI data, including a smaller peak change at the junction between the striate cortex and the cuneus in the right hemisphere (Howard et al., 1992) and in the central/precentral sulcus, cerebellar vermis, right thalamus, and anterior cingulate gyrus (Price et al., 1996). The present study did not reveal a general pattern of activation in any of these areas, although three subjects showed right cerebellar activity, which has previously been reported in studies involving language production, such as the auditory repetition compared to rest (Petersen et al., 1989) and verb generation (Desmond & Fiez, 1998).

CONCLUSIONS

Overall, whereas the results of the present study confirm the findings of previous PET studies, showing that *on average*, the posterior STG plays a specialized role in auditory processing of words, the study also demonstrates the extent of individual variability in the location of the activation along the STG. In addition, only some subjects showed IFG activation, perhaps explaining why activation in this region is often less robust (cf. Price et al., 1996). The group and individual results taken together suggest that although brain encodings for language may fall into regular patterns, and may be quite similar among individuals, they are different enough to complicate the interpretation of group average neuroimaging data. They also suggest that the apparent disparate results from single case lesion data and neuroimaging data might not be so different as sometimes declared.

Finally, a combination of careful head immobilization, instructions to the subject to minimize jaw movement, and careful post hoc movement correction and analysis permits the performance of fMRI studies during overt articulation. Further, the use of a formal measure to quantify the movement artifact in the oropharynx represents an objective rather than subjective criterion to exclude subjects with significant oropharyngeal motion. With these methods, subjects performing overt articulation show reliable activation in areas typically associated with lexical processing.

The present results are consistent with traditional neuropsychological models of auditory lexical processing and help explain the presence of individual differences that are common in studies of patients with brain lesions.

REFERENCES

- Bandettini, P. A., Jesmanowicz, A., Wong, E. C., & Hyde, J. S. (1993). Processing strategies for time-course data sets in functional MRI of the human brain. *Magnetic Resonance in Medicine*, *30*, 161–173.
- Barch, D. M., Sabb, F. W., Carter, C. S., Braver, T. S., Noll, D. C., & Cohen, J. D. (1999). Overt verbal responding during fMRI scanning: Empirical investigations of problems and potential solutions. *Neuroimage*, *10*, 642–657.
- Birn, R. M., Bandettini, P. A., Cox, R. W., & Shaker, R. (1999). Event-related fMRI of tasks involving brief motion. *Human Brain Mapping*, *7*, 106–114.
- Burton, M. W., & Small, S. L. (1999). An introduction to fMRI. *Neurologist*, *5*, 145–158.
- Caramazza, A. (1986). On drawing inferences about the structure of normal cognitive systems from the

- analysis of patterns of impaired performance: The case for single-patient studies. *Brain and Cognition*, **5**, 41–66.
- Chee, M. W., O'Craven, K. M., Bergida, R., Rosen, B. R., & Savoy, R. L. (1999). Auditory and visual word processing studied with fMRI. *Human Brain Mapping*, **7**, 15–28.
- Cohen, J. D., MacWhinney, B., Flatt, M. R., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments & Computers*, **25**, 257–271.
- Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology*, **33A**, 497–505.
- Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers in Biomedical Research*, **29**, 162–173.
- Dejerine, J. (1891). Sur un cas de cécité verbale avec agraphie, suivi d'autopsie. *Comptes Rendus Hebdomadaires des Séances et Mémoires de la Société de Biologie*, **3**, 197–201.
- Démonet, J. F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J.-L., Wise, R., Rascol, A. & Frackowiak, R. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain*, **115**, 1753–1768.
- Démonet, J. F., Fiez, J. A., Paulesu, E., Petersen, S. E., & Zatorre, R. J. (1996). PET studies of phonological processing: A critical reply to Poeppel. *Brain and Language*, **55**(3), 352–379.
- Desmond, J., & Fiez, J. (1998). Neuroimaging studies of the cerebellum: language, learning and memory. *Trends in Cognitive Sciences*, **2**, 355–362.
- Donders, F. C. (1868/1969). On the speed of mental processes. *Acta Psychologica*, **30**, 412–431.
- Dronkers, N. (1998). Symposium: The role of Broca's area in language. *Brain and Language*, **65**, 71–72.
- Fiez, J., Raichle, M., Balota, D., Tallal, P., & Petersen, S. (1996). PET activation of posterior temporal regions during auditory word presentation and verb generation. *Cerebral Cortex*, **6**, 1–10.
- Frackowiak, R. S. J., Friston, K. J., Frith, C. D., Dolan, R. J. & Mazziotta, J. C. (1997). *Human brain function*. New York: Academic Press.
- Francis, W. N., & Kucera, H. (1982). *Frequency analysis of American English*. Boston: Houghton-Mifflin.
- Friston, K. J., Price, C. J., Fletcher, P., Moore, C., Frackowiak, R. S. J., & Dolan, R. J. (1996). The trouble with cognitive subtraction. *Neuroimage*, **4**, 97–104.
- Gabrieli, J. D., Poldrack, R. A., & Desmond, J. E. (1998). The role of left prefrontal cortex in language and memory. *Proceedings of the National Academy of Sciences USA*, **95**, 906–913.
- Grady, C. L., Van Meter, J. W., Maisog, J. M., Pietrini, P., Krasuski, J., & Rauschecker, J. P. (1997). Attention-related modulation of activity in primary and secondary auditory cortex. *NeuroReport*, **8**, 2511–2516.
- Hajnal, J., Saeed, N., Soar, E., Oatridge, A., Young, R., & Bydder, G. (1995). A registration and interpolation procedure for subvoxel matching of serially acquired MR images. *Journal of Computer Assisted Tomography*, **19**, 289–296.
- Howard, D., Patterson, K., Wise, R., Brown, W. D., Friston, K., Weiller, C., & Frackowiak, R. (1992). The cortical localization of the lexicons: Positron emission tomography evidence. *Brain*, **115**, 1769–1782.
- Kinahan, P. E., & Noll, D. C. (1999). A direct comparison between whole-brain PET and BOLD fMRI measurements of single-subject activation response. *Neuroimage*, **9**, 430–438.
- Mesulam, M. (1990). Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Annals of Neurology*, **28**, 597–613.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Noll, D. C., Cohen, J. D., Meyer, C. H., & Schneider, W. (1995). Spiral K-space MRI imaging of cortical activation. *Journal of Magnetic Resonance Imaging*, **5**, 49–56.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, **9**, 97–113.
- O'Leary, D. S., Andreason, N. C., Hurtig, R. R., Hichwa, R. D., Watkins, G. L., Ponto, L. L., Rogers, M., & Kirchner, P. T. (1996). A positron emission tomography study of binaurally and dichotically presented stimuli: Effects of level of language and directed attention. *Brain and Language*, **53**, 20–39.
- Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M. A., & Raichle, M. E. (1989). Positron emission

- tomographic studies of the processing of single words. *Journal of Cognitive Neuroscience*, **1**, 153–170.
- Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M. A., & Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature*, **331**, 585–589.
- Poeppl, D. (1996). A critical review of PET studies of phonological processing. *Brain and Language*, **55**, 317–351.
- Price, C. J., Wise, R. J., Warburton, E. A., Moore, C. J., Howard, D., Patterson, K., Frackowiak, R. S., & Friston, K. J. (1996). Hearing and saying: The functional neuroanatomy of auditory word processing. *Brain*, **119**, 919–931.
- Pugh, K. R., Shaywitz, B. A., Shaywitz, S. E., Fulbright, R. K., Byrd, D., Skudlarski, P., Shankweiler, D., Katz, L., Constable, R. T., Fletcher, J., Lacadie, C., Marchione, K., & Gore, J. C. (1996). Auditory selective attention: An fMRI investigation. *Neuroimage*, **4**, 159–173.
- Rueckert, L., Appollonio, I., Grafman, J., Jezzard, P., Johnson, R., Jr., Le Bihan, D., & Turner, R. (1994). Magnetic resonance imaging functional activation of left frontal cortex during covert word production. *Journal of Neuroimaging*, **4**, 67–70.
- Saberi, K., & Perrott, D. R. (1999). Cognitive restoration of reversed speech. *Nature*, **398**, 760. [letter].
- Sergent, J., Zuck, E., Lévesque, M., & MacDonald, B. (1992). Positron emission tomography study of letter and object processing: Empirical findings and methodological considerations. *Cerebral Cortex*, **2**, 68–80.
- Small, S. L., Burton, M. W., Perfetti, C. A., & Noll, D. C. (1999). Listening to sentences and to sequential word strings. Manuscript submitted for publication.
- Small, S. L., Noll, D. C., Perfetti, C. A., Hlustik, P., Wellington, R., & Schneider, W. (1996). Localizing the lexicon for reading aloud: Replication of a PET study using fMRI. *NeuroReport*, **7**, 961–965.
- Steinmetz, H., & Seitz, R. J. (1991). Functional anatomy of language processing: Neuroimaging and the problem of individual variability. *Neuropsychologia*, **29**, 1149–1161.
- Sternberg, S. (1969). The discovery of processing stages: Extensions of the Donders' method. *Acta Psychologica*, **30**, 276–315.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain: 3D proportional system: An approach to cerebral imaging*. New York: George Thieme Verlag.
- Van Essen, D. C., Drury, H. A., Joshi, S., & Miller, M. I. (1998). Functional and structural mapping of human cerebral cortex: Solutions are in the surfaces. *Proceedings of the National Academy of Sciences USA*, **95**, 788–795.
- Wernicke, C. (1874). *Der Aphasische Symptomenkomplex*. Breslau: Cohn & Weigert.
- Wildgruber, D., Ackerman, H., Klose, U., Kardatzki, B., & Grodd, W. (1996). Functional lateralization of speech production at primary motor cortex: a fMRI study. *NeuroReport*, **7**, 2791–2795.
- Wise, R., Chollet, F., Hadar, U., Friston, K., Hoffner, E., & Frackowiak, R. (1991). Distribution of cortical neural networks involved in word comprehension and word retrieval. *Brain*, **114**, 1803–1817.
- Woods, R. P., Cherry, S. R., & Mazziotta, J. C. (1992). Rapid automated algorithm for aligning and reslicing PET Images. *Journal of Computer Assisted Tomography*, **16**, 620–633.
- Yetkin, F. Z., Hammeke, T. A., Swanson, S. J., Morris, G. L., Mueller, W. M., McAuliffe, T. L., & Houghton, V. M. (1995). A comparison of functional MR activation patterns during silent and audible language tasks. *American Journal of Neuroradiology*, **16**, 1087–1092.