PET Activation of Posterior Temporal Regions during Auditory Word Presentation and Verb Generation

Previous studies using positron emission tomography (PET) report blood flow changes in superior and middle temple gyri associated with auditory and language tasks (Petersen et al., 1988, 1989; Wise et al., 1991; Demonet et al., 1992; Howard et al., 1992; Sergent et al., 1992; Zatorre et al., 1992; Petrides et al., 1993; Raichle et al., 1994; Fiez et al., 1995). An important issue is whether these changes reflect the activation of a single functional region or multiple regions with distinct functional contributions. In the present study, we examined this issue by focusing upon two tasks for which we have previously reported posterior temporal blood flow changes: listening to auditorily presented words (Petersen et al., 1988, 1989), and generation of a verb in response to a visually presented noun (Raichle et al., 1994); see also Wise et al. (1991). We began by further characterizing a left temporoparietal region of change previously associated with auditory word presentation. This previously reported response was replicated, and the results were extended by demonstrating presentation of pseudowords also produced activation. We next asked whether the activation associated with auditory word presentation could be distinguished from that associated with the generation of verbs in response to visually presented nouns. It was found that the activations associated with these two tasks could be both functionally and spatially dissociated. Thus, two posterior temporal areas associated with auditory word presentation and verb generation appear to represent distinct areas concerned with word processing. More generally, the results demonstrate an approach for assessing the independence of two activated areas.

Blood flow changes in the posterior portions of the superior and middle temple gyri have been found in a number of different studies using positron emission tomography (PET) report blood flow changes in superior and middle temple gyri associated with auditory and language tasks (Petersen et al., 1988, 1989; Wise et al., 1991; Demonet et al., 1992; Howard et al., 1992; Raichle et al., 1994). While across these studies many of the tasks share similarities (e.g., auditory presentation of stimuli (Petersen et al., 1988, 1989; Wise et al., 1991; Demonet et al., 1992; Howard et al., 1992)), there are also significant differences between the tasks; for example, in some cases subjects merely listened to auditorily presented words (Petersen et al., 1988, 1989; Wise et al., 1991), while in other cases subjects performed semantic or phonological judgments upon presented words (Wise et al., 1991; Demonet et al., 1992).

At present, it is unclear whether the posterior temporal changes observed across different PET studies reflect the activation of a single functional region, or whether multiple regions with distinct functional contributions can be defined. In the present study, we examine this issue by focusing upon posterior temporal and temporoparietal blood-flow increases we have previously associated with the performance of two different language tasks. For one task, subjects listened to auditorily presented nouns while maintaining visual fixation on the small dot. In the control condition, subjects maintained fixation on a small dot. Bilateral activation was noted along the middle portions of the superior temporal gyrus, at or near primary and extraparyetal auditory areas (Brodmann areas 41 and 42). In addition, a left-lateralized temporoparietal focus of activation was found more posteriorly, at or near the posterior portion of Brodman area 22 (Wernicke's area).

In contrast to the primary and extraparyetal regions, the left temporoparietal region identified by Petersen et al. (1988, 1989) did not appear to be activated by passive presentation of relatively simple auditory stimuli, such as tone pips, clicks, and noise bursts (Roland et al., 1981; Mazziota et al., 1982; Lauter et al., 1985; Zatorre et al., 1992). On the basis of this pattern of activation, Petersen et al. (1988, 1989) hypothesized that the temporoparietal region might be related to speech-specific acoustic processing. Functional dissociations between linguistic and nonlinguistic stimuli have previously been reported in the visual modality (Petersen et al., 1990; Howard et al., 1992). For instance, Petersen et al. (1990) reported that visually presented words and pseudowords (pronounceable letter strings such as "floog") activated a group of contiguous areas in left medial extrastrate cortex, but not visually presented strings of consonant letters and letter-like forms (false fonts). In this study, we were interested in determining whether the left temporoparietal area activated by auditory presentation of words might also be activated by auditory presentation of pseudowords, in a manner analogous to the medial extrastrate activation observed for visual presentation of words and pseudowords, and as suggested by previous work in the auditory modality (Wise et al., 1991).

The experimental paradigm had the additional benefit of allowing us to determine whether the left temporoparietal activation associated with auditory word presentation could be replicated across experiments. This was of particular concern because of results found in a related study designed to investigate neural aspects of speech and temporal perception (Fiez et al., 1995). In the Fiez et al. study, subjects were presented with synthetic words under both a passive and an active detection condition. Contrary to our expectations, under these conditions significant left temporoparietal activation near that found by Petersen et al. (1988, 1989) was not found. This failure raised questions about the reliability of the activation noted by Petersen et al. (1988, 1989), and the possible importance of the task and stimulus differences between the studies by Fiez et al. (1995) and Petersen et al. (1988, 1989).

The second issue addressed in this study was the relationship between the left posterior temporal activation associated with the auditory presentation of words, and that associated with another task: generation of verbs in response to visually presented nouns (Raichle et al., 1994). In the study by Raichle et al., subjects were instructed to think of and say aloud appropriate verbs for visually presented nouns (what the nouns might be used for, or what they might do). As a control condition, subjects read aloud visually presented nouns. Greater activation in a posterior temporal region was observed when subjects performed the verb generation task than when the same subjects read aloud visually presented nouns (Raichle et al., 1994). Along with left frontal, right cerebellar, and anterior cingulate regions, the left temporal region became signifi-
in significantly less active following practice on the verb generation task (Raichle et al., 1994); the magnitude of the left temporal activation also appeared to be affected by the rate of stimulus presentation, with significantly greater activation observed when stimuli were presented at a rate of 1 every 1.5 sec (Raichle et al., 1994) than when they were presented at a rate of 1 per second (Petersen et al., 1988, 1989).

In the second section of the present report, our goal was to determine whether the regions of activation associated with the two tasks are both functionally and spatially distinct. This issue was of interest because although the two tasks differ significantly in their processing demands, some have hypothesized that they should share common functional regions of activation. For instance, Wise et al. (1991) theorized that both verb generation and listening to auditorily presented words involve, in part, a similar form of semantic analysis that accounts for the same posterior temporal activation found during both tasks. More generally, it has also been hypothesized that language processing involves a distributed set of regions that are automatically activated by presentation of linguistic material; thus, a common set of areas are expected to be activated during a wide range of language tasks (Debrenet et al., 1993).

Replication and Extension of Auditory Word Presentation

Materials and Methods

The first issue we addressed was the response properties of the left temporoparietal area activated by auditory presentation of words (Petersen et al., 1988, 1989). In particular, we were interested in determining whether this left temporoparietal area might also be activated by auditory presentation of pseudowords. The experimental paradigm had the additional benefit of allowing us to establish the reliability of responses in the left temporoparietal area across experiments.

Subjects

Subjects were normal volunteers recruited from the undergraduate and medical campuses of Washington University. All were strongly right-handed as judged by the Edinburgh handedness inventory (Raczkowski et al., 1974). Subjects were paid $75 for their participation and gave informed consent in accordance with guidelines set by the Human Studies and the Radioactive Drug Research Committees of Washington University. Data were collected from three subjects: two males (age 21 and 26 years), and one female, 24 years old.

Stimuli

Word stimuli consisted of familiar nouns, verbs, and adjectives. They ranged in visual frequency from 1 to 1360 occurrences per million words, with a mean of 122 ± 12 SE occurrences per million words (Kucera and Francis, 1967). The words were divided into three lists of 55 items. For each list of words a corresponding list of pseudowords was created by substituting one or more phonemes in each word to create a pseudoword (e.g., the matched pseudoword for the word "yard" was "dard"). Stimuli were recorded onto audio tape and presented at the rate of one item per second. The spoken duration of the words ranged from 330 to 785 msec, with a mean duration of 548 ± 6 msec SE. The spoken duration of the pseudowords ranged from 327 to 801 msec, with a mean duration of 566 ± 6 msec SE.

The tape-recorded stimuli were replayed using an audio tape deck (Lux Corporation, Luxman model K-111). Audio signals from the tape deck and computer sound channel were amplified (Pioneer Electronic Corporation, model SX-1700), and retransmitted binaurally through receivers (Knowles insert receiver, model 1912) especially designed to snap into fitted plastic earmolds placed in each subject's ears (Lateur et al., 1985). The system was designed to mimic the filter characteristics of the human pinna and outer ear (Lateur et al., 1985).

Paradigm

Nine scans were conducted in each subject. Scans 2, 5, and 8 were control scans in which subjects were instructed to maintain fixation on a 5 mm dot displayed on a monitor suspended approximately 15 cm in front of them. For the active scans (scans 1, 3, 4, 6, 7, and 9), subjects again maintained fixation on the displayed dot, but also were instructed to listen to the presented words or pseudowords. Each of the six lists was presented during a separate active scan, alternating between word and pseudoword lists. The order of list presentation (e.g., word and matched pseudoword lists A, B, and C) varied across subjects.

Image Acquisition

Each subject laid on the scanner couch and a venous catheter was placed in the right arm. Subjects' heads were immobilized within a closely fitted, thermally molded, plastic facial mask individually made for each subject. Head alignment and planes of section were recorded with a lateral skull x-ray (Fox et al., 1985). For each scan, 18O-labeled carbon dioxide (C18O2) was bubbled through saline to form 18O-labeled water (Welch and Kilbourn, 1985), which was administered as an intravenous bolus of 8-10 ml of saline containing 1 mCi/kg of body weight. Data were acquired using the PETT VI system employed in the low-resolution mode, which simultaneously acquires seven parallel slices with a center-to-center distance of 14.4 mm (Ter-Pogossian et al., 1982; Yamamoto et al., 1982). Images were reconstructed using filtered backprojection to a resolution of 17 mm full width at half-maximum.

Since blood-flow and radiation counts are nearly linearly related over normal physiological ranges, images of radioactive counts rather than blood flow were utilized. This eliminated the risk and discomfort of arterial catherization (Herscovitch et al., 1983; Fox and Mintun, 1989). In order to negate the effects of global fluctuations in blood flow and variations in the artificial arterial input function, the number of counts in each image was normalized to 1000 (Fox et al., 1987); magnitude data will therefore be given in terms of normalized counts.

Image Transformations

To place subjects into a standard anatomical format that allows comparisons to be made across subjects, each image was also transformed into the space of the Talairach stereotactic atlas (Fox et al., 1985, 1988; Talairach and Tournoux, 1988). Foci of change will be reported by listing three coordinates (x, y, z): the x-coordinate is the distance in mm to the left (−) or right (+) of midline, the y-coordinate is the distance in mm anterior (+) or posterior (−) to the anterior commissure, and the z-coordinate is the distance in mm above (+) or below (−) a horizontal plane which cuts through the anterior-posterior commissures.

The basis of all data analysis was a comparison of neuronal activity under two conditions. To isolate areas of activity change between task conditions, images generated during performance of one task were subtracted, on a pixel-by-pixel basis, from images generated during performance of another task. Individual difference images were created by subtracting the images acquired during different two scans (an "active task scan" and a "control scan") performed in a single subject. Average difference images were created by adding individual difference images together and then dividing the resulting images by the number of contributing individual images; this yields an image of the mean regional changes across subjects and/or task conditions (Fox et al., 1988).

Data Analysis

Analysis of Temporoparietal Regional Magnitudes. A region of interest approach was used to assess whether the temporoparietal activation associated with auditory word presentation (Petersen et al., 1988, 1989) could be replicated across subjects, and whether the finding could be extended to auditory presentation of pseudowords. A temporoparietal region of interest was first identified using a passive auditory word minus fixation image created from a separate group of eight subjects (Petersen et al., 1988, 1989). The location of the peak temporoparietal response in this previous study was reported in terms of the Talairach 1967 atlas (Talairach et al., 1967), at x = −54, y = −30, z = 14. To define this response in terms of the Talairach and Tournoux (1988) atlas, data from the study by Petersen et al. (1988, 1989) were cast into the space of the Talairach 1988 atlas (as described in the Materials and Methods). Next, the location of the peak temporoparietal activation in this image was used to define a spherical region with a 7 mm radius centered at x = −53, y = −49, z = 18. The 7 mm radius was chosen because this value corresponds to the resolution of the algorithm used to automatically
identify all regions of change in individual and average difference images (Mintun et al., 1989).

Next, the region of interest was placed in each of the nine passive auditory word minus fixation and the nine auditory pseudoword minus fixation individual difference images from the present study (three repetitions of each stimulus condition in three subjects). The regional magnitude in each individual difference image was determined by calculating the average magnitude of all pixels located inside of the region of interest.

The 18 individual regional magnitudes were first analyzed using a two-factor ANOVA to determine whether significant differences between subjects or task repetitions existed. No significant effects of task repetition \(F(2, 3) = 0.60, p = 0.68\) or subject \(F(2, 2.44) = 0.14\) were found. To determine whether the regional temporoparietal activation was significantly greater than zero, as hypothesized on the basis of previous results (Petersen et al., 1988, 1989), the regional magnitudes from the nine auditory word minus fixation difference images were then analyzed using a one-sample \(t\) test. In order to determine whether the temporoparietal activation extended to pseudowords, the same analysis was performed using the nine passive auditory pseudoword minus fixation individual difference images. A paired \(t\) test was used to compare the magnitude of regional activation in the words and pseudowords conditions.

Analysis of Temporoparietal Activation Locations. One limitation to a regional analysis is that a significant change might be found even though the region is encompassing only a portion of some large response centered fairly distantly. To compare the locations of temporoparietal activation between our earlier work (Petersen et al., 1988, 1989) and the current study, auditory stimulation minus fixation images from the present study were averaged across subjects and repetitions of the task conditions, yielding an \(N = 9\) average difference image for each stimulus type (three subjects \(\times\) three repetitions each of words and pseudowords). In each average difference image, the location and peak magnitude of the response focus nearest to the region of interest defined from our previous work \((x = -53, y = -49, z = 18)\) was identified.

Identification of Other Regions of Activation. Though this study was motivated by a priori regions of interest from our previous work, we also examined the auditory word minus fixation and pseudoword minus fixation average difference images for other regions of activation using a change-distribution analysis (Fox et al., 1988, 1989; Mintun et al., 1989). For both the auditory word minus fixation and the pseudoword minus fixation average difference images, all regions of change, both positive and negative, were first identified (Mintun et al., 1989). The distribution of foci magnitudes from each image was then compared to a standard noise distribution in order to determine whether the distribution was positively and/or negatively leptokurtotic, that is, whether there were significant positive or negative outliers in the distribution. If there were significant outliers, post hoc \(t\) values and corresponding \(p\) values were computed for each focus (Fox et al., 1988; Fox and Mintun, 1989). Because of limitations of this and other thresholding techniques (Hunton et al., 1994), the results from the outlier analysis should be considered to be exploratory, or hypothesis-generating, examination of the data.

Results

In total, the results described below provide a clear replication of activation associated with word presentation relative to Petersen et al. (1988, 1989), and extend the findings to pseudowords.

Magnitude Analysis

There were significant changes of activity at the temporoparietal region of interest for both words and pseudowords (see Fig. 1). The mean magnitude of activation (in counts) across the individual difference images was \(38 \pm 13\) SE for auditory presentation of words, and \(34 \pm 11\) SE for pseudowords. Both values are significantly more positive than zero \(t(8) = 3.17, p = 0.01\), and \(t(8) = 2.87, p = 0.007\), respectively. This regional activation did not significantly differ between words and pseudowords \(t(8) = 0.22, p = 0.83\).

Location Analysis

Results from the location analysis provide strong evidence the change measured within the temporoparietal region reflects the activity of a region that is distinct from the nonauditory regions. First, the posterior temporal words minus fixation and pseudowords minus fixation average difference images, temporoparietal increases in activity were found very near the temporoparietal response to passive auditory words we previously reported (Petersen et al., 1988, 1989) at \(x = -53, y = -49, z = 18\). For words, the center of the increase was located within \(7\) mm, at \(x = -51, y = -55, z = 18\), with a peak magnitude of \(49\) counts. For pseudowords, the center of the increase was also within \(7\) mm, at \(x = -51, y = -55, z = 20\), with a peak magnitude of \(46\) counts.

Identification of Other Regions of Activation

Using an outlier analysis to identify significant regions of change, significant leptokurtosis was found only for the positive distribution of foci magnitudes in both the auditory word minus fixation and the auditory pseudoword minus fixation average difference images. A post hoc analysis was used to identify all positive foci of change with \(Z\) scores corresponding to a \(p < 0.01\). In the auditory word minus fixation image this identified one focus located at or near Brodmann area 41 on the right \((x = 55, y = -15, z = 8)\) and two foci on the left \((x = -51, y = -21, z = 14\), and \(x = -59, y = -29, z = 16)\). In the auditory pseudoword minus fixation similar results were found, with two foci found at or near Brodmann areas 41 and 42 on the right \((x = 53, y = -15, z = 10)\) and one on the left \((x = -53, y = -21, 12)\). These results are consistent with the primary and secondary auditory responses we have previously associated with auditory word presentation (Petersen et al., 1988, 1989). Neither of the temporoparietal peaks identified in the location analysis reached significance using the outlier analysis; this failure is not unexpected, given recent evidence that the outlier analysis and other thresholding techniques are fairly conservative approaches which may fail to detect many reliable regions of change (Hunton et al., 1994).

Differences between Verb Generation and Auditory Tasks

Materials and Methods

The second issue we addressed was the relationship between the left temporoparietal area (Petersen et al., 1988, 1989) and the more ventrally situated posterior temporal area, which became active during the generation of verbs to visually presented nouns (Rauchle et al., 1994). The activation of these two areas was examined both functionally and spatially in order to determine whether, within a confined area of left posterior temporal cortex, distinct areas concerned with processing words can be identified with PET functional imaging studies. Both analyses were based upon first defining regions of posterior temporal activation associated with each task condition. Next, the relationship between task condition (passive listening or verb generation) and the magnitude and location of activation relative to the two regions of interest was examined in independent groups of subjects.

Region of Interest Generation

We first defined posterior temporal regions of interest associated with both the passive listening and the verb generation task conditions. This was accomplished using data from one group of subjects who performed the passive listening task, and data from another group of subjects who performed the verb generation task.

To generate a region of interest using data from subjects listening to auditory words, an \(N = 9\) average difference image was first created from the word presentation minus fixation images from the present study (three subjects \(\times\) three repetitions). In this average image, the largest response that localized to left temporal cortex [with the exception of primary and extraparital auditory cortex, as...