Auditory Hallucinations and the Temporal Cortical Response to Speech in Schizophrenia: A Functional Magnetic Resonance Imaging Study


Objective: The authors explored whether abnormal functional lateralization of temporal cortical language areas in schizophrenia was associated with a predisposition to auditory hallucinations and whether the auditory hallucinatory state would reduce the temporal cortical response to external speech.

Method: Functional magnetic resonance imaging was used to measure the blood-oxygenation-level-dependent signal induced by auditory perception of speech in three groups of male subjects: eight schizophrenic patients with a history of auditory hallucinations (trait-positive), none of whom was currently hallucinating; seven schizophrenic patients without such a history (trait-negative); and eight healthy volunteers. Seven schizophrenic patients were also examined while they were actually experiencing severe auditory verbal hallucinations and again after their hallucinations had diminished.

Results: Voxel-by-voxel comparison of the median power of subjects’ responses to periodic external speech revealed that this measure was reduced in the left superior temporal gyrus but increased in the right middle temporal gyrus in the combined schizophrenic groups relative to the healthy comparison group. Comparison of the trait-positive and trait-negative patients revealed no clear difference in the power of temporal cortical activation. Comparison of patients when experiencing severe hallucinations and when hallucinations were mild revealed reduced responsivity of the temporal cortex, especially the right middle temporal gyrus, to external speech during the former state.

Conclusions: These results suggest that schizophrenia is associated with a reduced left and increased right temporal cortical response to auditory perception of speech, with little distinction between patients who differ in their vulnerability to hallucinations. The auditory hallucinatory state is associated with reduced activity in temporal cortical regions that overlap with those that normally process external speech, possibly because of competition for common neurophysiological resources.

(Am J Psychiatry 1997; 154:1676–1682)
vials, such as reduced right ear (left hemisphere) advantage on dichotic listening tasks (10) and reduced left temporal cortical activation when listening to single words (11) or when monitoring inner speech.

Auditory hallucinations are a core feature of schizophrenia (12). Neuroimaging studies using positron emission tomography (13–16) and single photon emission tomography (17, 18) indicate that auditory association and language cortices are active in schizophrenic patients with auditory hallucinations. However, only a minority of these studies (16–18) have examined brain activity during the hallucinations. Moreover, most studies have limited their measurements to predefined regions, often within a single brain slice, rather than examining activity across the whole temporal cortex. Silbersweig et al. (16), while avoiding this problem, found that temporal cortical activity coincident with hallucinations in individual scans did not reach accepted levels of significance on averaged group images. This, the authors argued, may have been due to intersubject variation in the functional anatomy of the temporal cortex.

Functional magnetic resonance imaging (MRI) is sensitive to changes in cerebral blood oxygenation related to neural activity. The lack of a requirement for administration of radioactive isotopes makes functional MRI particularly suitable for studies of patients. We have previously used functional MRI to show that in a single schizophrenic patient, auditory hallucinations were associated with activation in the right middle temporal gyrus, one of the regions that normally responds to external speech (19). In the present study, we examined the functional anatomy of the cortical response to speech in schizophrenic subjects with a history of hallucinations, schizophrenic subjects without a history of hallucinations, and healthy comparison subjects. We predicted that the patients with a history of hallucinations would exhibit a less left-lateralized response to auditory perception of externally presented speech than both the healthy comparison subjects and the patients without a history of hallucinations. We also predicted that auditory hallucinations would specifically reduce the responsivity of the right temporal cortex to external speech. We tested this hypothesis by comparing severely hallucinating patients before and after improvement.

**METHOD**

Functional MRI data on eight right-handed (20) male schizophrenic patients (DSM-III-R diagnosis) who had a history of auditory hallucinations but were not actively hallucinating (trait-positive) were compared with eight healthy male subjects recruited from the Institute of Psychiatry (mean age=35.3 years, SD=6.3) and seven male schizophrenic patients (DSM-III-R diagnosis) who had not experienced auditory hallucinations (trait-negative, as defined by McGuire et al. [21]) (table 1). Clinical ratings of the patients were made with the Scale for the Assessment of Positive Symptoms (SAPS) (22) and the Scale for the Assessment of Negative Symptoms (SANS) (23); these were recorded on the day of scanning. The normal comparison subjects were not taking medications and had neither smoked nor taken alcohol on the morning before being scanned.

A further group of seven right-handed male schizophrenic subjects underwent functional MRI studies on two separate consecutive occasions (1) during a period of severe ongoing auditory verbal hallucinations (hallucinatory-state-positive) and (2) after these had diminished (hallucinatory-state-negative) (table 1). Four patients were common to both the group that was studied during severe and less severe hallucinatory states and the trait-positive group. The mean interval between the two scanning sessions was 80.3 days (SD=9.4, range=34–118). The severity of auditory hallucinations was assessed on the day of each scanning session. Subjective ratings of hallucination severity (0=no hallucinations; 10=worst ever experienced) were obtained for any hallucinations that occurred immediately before or during scanning.

After complete description of the study to the subjects, written informed consent was obtained.

**Image Acquisition**

At each of 10 noncontiguous 5-mm-thick oblique slices through fronto-temporo-occipital regions, 100 T_2*-weighted echoplanar magnetic resonance images depicting blood-oxygenation-level-dependent contrast (TE=40 msec, TR=3,000 msec) were acquired with
The use of a 1.5-T GE Signa system (retrofitted with an advanced nuclear magnetic resonance operating console). In the same scanning session, 43 3-mm-thick inversion recovery echoplanar anatomic images were acquired parallel to the anterior-posterior commissure line (TE=80 msec, TI=180 msec, TR=16 seconds, in-plane resolution=3 mm, eight signal averages).

Each 5-minute experiment followed a repeating A-B design. Auditory stimulation consisted of an identical prerecorded continuous

FIGURE 1. Brain Activation Maps

*Each row shows four axial slices through the auditory cortex from the middle temporal gyrus (left) to the posterior superior temporal gyrus (right); Talairach z coordinates in millimeters (inferior-superior) are shown below the bottom row. Each map is oriented with the right side of the brain represented on the left side of the map.

Top row: Generic brain activation maps separately computed for eight healthy comparison subjects and 15 schizophrenic patients with a history of auditory hallucinations (trait-positive) and without (trait-negative) are superimposed. Blue indicates brain regions generically activated in both groups (339 voxels), yellow indicates brain regions generically activated within the comparison group only (349 voxels), and red indicates regions activated in the schizophrenic group only (430 voxels). The comparison subjects activated the left temporal cortex more extensively than the schizophrenic patients, whereas the patients activated the right temporal cortex more extensively than the comparison subjects. The voxelwise probability of false positive activation in each generic brain activation map was $p < 0.0005$, uncorrected (expected number of false positive tests per map=10).

Second row: Maps show brain regions that demonstrated a significant difference in median power of response between the comparison group and the schizophrenic patients (trait-positive and trait-negative). Yellow indicates brain regions more powerfully activated in the comparison group (mainly located in the left temporal cortex) (29 voxels); red indicates brain regions more powerfully activated in the schizophrenic group (located in the right temporal cortex) (14 voxels). The voxelwise probability of a false positive difference between groups was $p<0.005$, uncorrected (search volume=1,029 voxels; expected number of false positive tests per map=5).

Third row: Generic brain activation maps separately computed for seven individuals with severe hallucinations (state-positive) and when hallucinations had diminished (state-negative) have been superimposed. Blue indicates brain regions generically activated at both times (109 voxels), yellow indicates brain regions activated only in the hallucinatory-state-positive group (122 voxels), and red indicates regions activated only in the hallucinatory-state-negative group (506 voxels). There was more extensive activation of the temporal cortex bilaterally in the state-negative group. The voxelwise probability of false positive activation in each generic brain activation map was $p<0.0005$, uncorrected (error pixels per image=10; expected number of false positive tests per map=10).

Fourth row: Maps show brain regions that demonstrated a significant difference in median power of response between the state-positive and state-negative conditions. Yellow indicates brain regions more powerfully activated in the state-positive condition (seven voxels); red indicates brain regions more powerfully activated in the state-negative condition (59 voxels, mostly in the right temporal cortex). The voxelwise probability of a false positive difference was $p<0.005$, uncorrected (search volume=737 voxels; expected number of false positive tests per map=4).
story presented binaurally through headphones for epochs of 39 seconds, alternating with 39 seconds of blank tape. Scanner noise was constant throughout the experiment. All subjects were instructed to keep their eyes closed and lie still. Head movement was restricted by using padding and a strap across the forehead.

**Image Analysis**

Each individual data set was corrected for the effects of the subject’s motion during image acquisition by means of a two-step procedure: realignment of the images by tricubic spline interpolation followed by regression of each realigned functional MRI time series on a second-order polynomial function of estimated positional displacements at that voxel across the course of the experiment (24). The power of periodic signal change at the fundamental frequency of alternation between the A and B conditions was estimated by fitting a sinusoidal regression model to the motion-corrected time series at each voxel of all images. The model comprised three pairs of sine and cosine waves at the fundamental frequency and the first two harmonic frequencies. One advantage of this model over some alternative techniques for estimating a periodic experimental effect in functional MRI time series, such as cross-correlation with the square wave or boxcar input function, is that it accommodates voxelwise variability in hemodynamic delay and waveform of response (25). The model is fitted by a pseudogeneralized least squares method (26, 27); that is, the model is initially fitted by ordinary least squares, a first-order autoregressive process is fitted to the ordinary least squares residual series, the estimated first-order autoregressive coefficient is used to transform the terms of the original model, and the transformed model is fitted by ordinary least squares. This iterative fitting procedure is necessary to ensure that the residuals ultimately satisfy criteria of independence and normality and that the estimates of model parameters and their standard errors are valid. At each voxel, the sum of squared amplitudes of the sine and cosine waves at the fundamental frequency yields an estimate of experimentally determined power in the time series, and this is divided by its standard error to yield an estimate of the fundamental power quotient. The fundamental power quotient is represented at each voxel by a parametric map. The observed time series are then randomly permuted 10 times, and the fundamental power quotient is estimated at each voxel after each permutation, exactly as described above. This results in 10 parametric maps of randomized fundamental power quotients at each anatomical plane of each individual data set (28).

To identify voxels that were activated “on average” within each group, the observed and randomized fundamental power quotient maps computed for each individual were registered in standard space (29) and identically smoothed with a Gaussian filter (full width at half maximum=7 mm) to accommodate individual variability in functional anatomy and error in registration. The registered procynthetic input. 1) finding the three-dimensional translations, rotations, and linear rescaling factors that minimized the difference in total gray scale intensity between each functional echoplanar image data set and the same individual’s structural echoplanar image data set with the use of the Fletcher-Davidon-Powell multidimensional search algorithm (30); 2) registering each individual’s structural echoplanar image data set with a histogram-matched template image in standard space by affine transformation; and 3) serially applying the transformation vectors defining steps 1 and 2 to each (observed or randomized) fundamental power quotient map. We then computed the observed median fundamental power quotient at each voxel where the images overlapped sufficiently and compared this with a randomized distribution of median fundamental power quotients ascertained from the randomized fundamental power quotient maps. Because we had previously permuted each time series 10 times to create the randomized fundamental power quotient maps, we were able to sample at least 200,000 estimates of the median fundamental power under the null hypothesis over a search volume of at least 20,000 voxels. We subsequently determined that this size of randomization distribution is sufficient to ascertain stable critical values for significance tests with p values of 0.0001 or greater (31, 32). The null hypothesis of no generic response to the experimental design was refuted at a given voxel if the observed median fundamental power quotient exceeded the critical value of the randomization distribution corresponding to a one-tailed test with a p value of 0.0005. Such generally activated voxels were colored and overlaid on the gray scale intensity of the template image to form a generic brain activation map, and regions of activation were identified by reference to the brain atlas of Talairach and Tournoux (24, 29).

To identify voxels that demonstrated a significantly different power of periodic response between groups, the following procedures were adopted. The median fundamental power quotient observed at each voxel was estimated separately for each group (as above), and the observed between-group difference in median fundamental power quotient was computed and represented in a parametric map. Individual subjects were then randomly reassigned to one or both of the two groups, and the between-group difference in median fundamental power quotient at each voxel was computed after each random reassignment. For the comparison of the hallucinatory-state-negative and -positive groups, individuals were reassigned in such a way that if the hallucinatory-state-positive image of a given individual was randomly assigned to one group, the hallucinatory-state-negative image of that individual was assigned to the other group. Random reassignment was repeated 10 times at each voxel, and differences in median fundamental power quotient after each random reassignment were pooled to sample the randomization distribution over the whole image. The observed between-group difference in median fundamental power quotient at the voxels that were generically activated within one or both of the groups was then compared to critical values of the randomization distribution corresponding to a two-tailed test with a p value of 0.005. Voxels at which the null hypothesis was refuted were colored and overlaid on the gray scale template image.

**RESULTS**

**Trait-Dermined Response to Speech**

All patients and comparison subjects in the trait-positive/trait-negative part of the study were Caucasian except for two Afro-Caribbean patients in the trait-negative group. The mean age, mean antipsychotic medication dosage, and mean clinical ratings of the trait-positive and trait-negative patient groups on the day of scanning are given in table 1. In addition to antipsychotic medication, three trait-positive patients were taking anticholinergics, one trait-positive and one trait-negative patient were taking a serotonergic antidepressant, one trait-positive patient was taking a tricyclic antidepressant, and another trait-positive patient was taking lithium.

Relative to the healthy comparison subjects, the schizophrenic patients (trait-positive and trait-negative groups combined) showed significantly less powerful and extensive activation of the left superior temporal gyrus (Brodmann’s area 22) (p<0.005, N=26 voxels, randomization test) and the auditory association cortex (Brodmann’s area 42) (p<0.005, N=3 voxels, randomization test) but more powerful and extensive response of the right middle temporal gyrus (Brodmann’s area 21) (p<0.005, N=10 voxels, randomization test) (figure 1, top and second rows). The median power of response of the comparison subjects significantly exceeded that of the schizophrenic patients in the left superior temporal gyrus (Talairach coordinates [mm]: x=−52, y=−17, z=−4; maximum between-group difference in maximum median fundamental power quotient=3.3) (p<0.005, Am J Psychiatry 154:12, December 1997 1679
randomization test) and in the left auditory association cortex (x = 61, y = 17, z = 9; maximum between-group difference in maximum median fundamental power quotient = 2.7) (p < 0.005, randomization test). The median power of response of the schizophrenic patients exceeded that of the healthy comparison subjects in the right middle temporal gyrus (x = 61, y = -8, z = 2; maximum between-group difference in maximum median fundamental power quotient = 1.2) (p < 0.005, randomization test).

There were minimal overall between-group differences in temporal cortical activations between the trait-positive and trait-negative groups. Six voxels bilaterally distributed in the temporal cortex demonstrated a more powerful response in the trait-positive than in the trait-negative patients; 11 voxels demonstrated a more powerful response in the trait-negative than in the trait-positive patients. The total number of differentially activated voxels (N = 17) was not much greater than the number of false positives (N = 4) expected for a test with a p value less than 0.005 and a search volume of 760 voxels.

State-Dependent Response to Speech

All patients in the state-positive/state-negative part of the study were Caucasian. Their mean age, mean antipsychotic medication dosage, and clinical ratings in the hallucinatory-state-positive and hallucinatory-state-negative conditions on the days of the two scanning sessions are given in table 1. The means of the two hallucination severity ratings (scale of 0–10) immediately before and during scanning in the two sessions (state-positive and state-negative) were 5.80 (SD = 2.0) and 1.64 (SD = 1.8), respectively (t = 4.10, df = 12, p < 0.001). Medication other than antipsychotics taken by the patients included anticholinergics (two state-positive subjects and one state-negative subject), serotoninergic antidepressants (two state-positive subjects and one state-negative subject), lithium (one state-positive and state-negative subject), diazepam and sodium valproate (one state-positive subject), and ranitidine (one state-positive subject).

External speech activated the temporal cortex significantly more powerfully and extensively in the hallucinatory-state-negative condition than in the hallucinatory-state-positive condition (figure 1, third and fourth rows). The greatest between-group differences in response were in the right middle temporal gyrus (Brodman’s area 21; N = 24 voxels) (Talairach x = 61, y = 6, z = 4; maximum between-group difference in maximum median fundamental power quotient = 1.2) (p < 0.005, randomization test) and in the left superior temporal gyrus (Brodman’s area 22; N = 20 voxels) (x = 58, y = -6, z = 4; maximum between-group difference in maximum median fundamental power quotient = 1.2) (p < 0.005, randomization test). There were seven voxels (from a total of 66 voxels) in which the power of response in the state-negative patients exceeded that in the state-positive patients.

DISCUSSION

Our conclusions are confined to measurements of temporal cortical activations because of the limited spatial coverage.

The combined groups of schizophrenic patients with and without a history of hallucinations showed less left-sided auditory cortical response to speech than the healthy comparison subjects. This relative hyporesponsivity to speech is consistent with evidence that schizophrenic patients have smaller left-sided superior temporal gyrus volumes than healthy comparison subjects.

In addition, external speech activated more right temporal cortex volume in the schizophrenic patients than in the healthy comparison subjects. This could be interpreted as an intrinsic hyperresponsivity of the right temporal cortex, if it were performing linguistic processing to a greater extent than it did in the healthy comparison subjects. Alternatively, a compensatory increase in right temporal cortical activity could be due to dysfunction in the left. Both interpretations are consistent with theories of normal lateralized brain development (6) and hypotheses of early damage to left hemisphere structures in schizophrenia. In the context of the finding by Bartel et al. (5) of significantly greater than normal right planum temporale area (but not volume) in schizophrenic patients, our results support the testable hypothesis that a right-sided shift in activation in this region in schizophrenia is related to underlying surface area rather than volume.

The functional implications of reversed laterality in schizophrenia have been hitherto largely unexplored. If, as has been proposed by Crow (6), this phenomenon has its origins in early brain development, it might be possible to detect abnormally lateralized auditory processes in children who will later develop schizophrenia.

No observable difference in either spatial extent or power of response between the trait-positive and trait-negative patients was found. In view of the presence of significant group differences in SANS scores and antipsychotic medication dosage, we cannot be certain that these factors did not obscure true group differences in temporal cortical activation.

Our results indicate that there is reduced response of the temporal cortex to external speech, particularly in the right middle temporal gyrus, during periods of increased auditory hallucinations. The intrasubject design overcomes many potential confounding factors that complicate functional neuroimaging research, such as intersubject variations in brain anatomy. The within-group comparison also enabled the detection of significant group differences in cortical response at lower maximum fundamental power quotient values than was possible for the intersubject comparisons.

The present findings are consistent with the hypothesis that auditory hallucinations “compete” with external speech for processing sites within the temporal cortex (19, 34, 35). This notion of competition is consistent with the use of listening to music or speech as a means of alleviating auditory hallucinations (12, 36). The narrative of a story...
may invoke more complex cognitive processes than those required to process the single words or simple sentences of auditory hallucinations. It is likely, therefore, that any competition between the two sets of stimuli for neurophysiological resources would be restricted to auditory cortex that subserves processing of the shared, simpler elements of speech.

Because we have not shown activity during auditory hallucinations, we cannot verify that this competition, if it exists, coincides exactly with the hallucinations themselves. Reduced auditory cortical response in the positive hallucinatory state might occur if hallucinations occurred only, or with greater frequency, during the nonauditory task intervals rather than during the listening intervals of the A-B design; however, no patients reported this when debriefed after scanning.

It is possible that the relative right-sided bias in modulation of the temporal cortex between the two hallucinatory states reflects the underlying functional anatomy of the schizophrenic patients who were also subjects in the trait-dependent part of the study. We can only infer that auditory hallucinations activate the right temporal cortex in these patients. However, this idea would be consistent with somestudies (13) that have demonstrated a predominance of right-sided temporal lobe activation in schizophrenic patients with auditory hallucinations, although most studies report bilateral or left-sided activity (15–18). This might reflect a selection bias, since many studies (e.g., MCGuire et al. [17]) restricted their analysis to the left side. Alternatively, it may be that hallucinatory activity competes with normal auditory processing only in regions responsible for processing both external speech and the hallucinatory experience.

Hallucinating schizophrenic patients are sometimes able to modulate the voices they hear by shifting attention to or away from them (12), and the greater the meaning of the stimuli (and attention to it), the more effective is the reduction in auditory hallucinations (36). Recent electrophysiological and neuroimaging evidence suggests that these attentional processes modulate activity at the auditory sensory cortical level (37, 38). Directing attention to an unexpected auditory stimulus leads to a greater temporal cortical response on the right than on the left (38–40). Thus, auditory hallucinations, which are usually unexpected, may modulate the response to speech more in the right temporal cortex than in the left. The right laterality of the effect might also be related to the emotional and prosodic qualities of auditory hallucinations, which are usually derogatory in content and tone (12)—features of speech associated with activity in the right temporal lobe (39). Future experiments that link sensory cortical activity to symptom-relieving attentional strategies could form the basis for assessing specific therapeutic interventions for hallucinations.

**CONCLUSIONS**

The normal response (left side greater than right) to external speech was reversed in schizophrenic patients (reduced on the left and increased on the right), but this effect did not distinguish those predisposed to auditory hallucinations from those not so predisposed. Periods of intense auditory hallucinations were associated with reduced power and spatial extent of temporal cortical response to external speech, particularly in the right middle temporal gyrus.


28. Andreasen NC: Scale for the Assessment of Positive Symptoms (SAPS). Iowa City, University of Iowa, 1984


