Functional Magnetic Resonance Imaging of Inner Speech in Schizophrenia

Claudia J.P. Simons, Derek K. Tracy, Kirandeep K. Sanghera, Owen O'Daly, James Gilleen, Maria-de-Gracia Dominguez, Lydia Krabbendam, and Sukhwinder S. Shergill

Background: Auditory verbal hallucinations in schizophrenia have been linked to defective monitoring of one's own verbal thoughts. Previous studies have shown that patients with auditory verbal hallucinations show attenuated activation of brain regions involved with auditory processing during the monitoring of inner speech. However, there are no functional magnetic resonance imaging studies explicitly comparing the perception of external speech with internal speech in the same patients with schizophrenia. The present study investigated the functional neuroanatomy of inner and external speech in both patients with schizophrenia and healthy control subjects.

Methods: Fifteen patients with schizophrenia and 12 healthy control subjects were studied using functional magnetic resonance imaging while listening to sentences or imagining sentences.

Results: Significant interactions between group (control subjects vs. patients) and task (listening vs. inner speech) were seen for the left superior temporal gyrus, as well as regions within the cingulate gyrus.

Conclusions: Attenuated deactivation of the left superior temporal gyrus in schizophrenia patients during the processing of inner speech may reflect deficits in the forward models subserving self-monitoring.

Key Words: Auditory verbal hallucinations, fMRI, inner speech, schizophrenia, temporal lobe, verbal self-monitoring

Auditory verbal hallucinations (AVHs) are one of the most common symptoms in schizophrenia. Current cognitive models suggest that auditory hallucinations are the result of defective self-monitoring (1–4). A feed-forward mechanism (5) has been proposed to explain self-monitoring of motor actions. In this model, motor commands that are needed to achieve a specific goal are identified and subsequently issued. Simultaneously, an efference copy is generated and transmitted through a corollary discharge mechanism to the sensory brain regions that are relevant to the planned act. The efference copy of the motor command serves to predict the sensory effects of the motor act. If the actual and predicted sensory feedback match, the actual sensory feedback will be attenuated. If a movement is externally controlled, there will be no efference copy that can attenuate the sensory information. Hence, the monitoring of motor acts can help to recognize whether a movement is self-initiated or externally generated. If generation of language and thought can be considered as a type of motor act (6–8), then monitoring of the verbal act may contribute to the distinction between self as source versus others as source. If this monitoring system is defective, verbal thoughts will not be recognized as being self-generated, leading to the experience of AVHs (8).

Neuroimaging studies investigating the generation of speech in healthy control subjects indicate that the monitoring process is associated with activation of the left prefrontal cortex and deactivation of the left and right superior temporal cortices (9), suggesting that corollary discharge from the areas involved in the generation of verbal speech results in attenuation of the activity in areas involved in speech perception. Initial studies of verbal fluency in schizophrenia indicated that this temporal deactivation was absent (9). However, more recent studies have failed to replicate this finding (10,11). Neuroimaging studies trying to "capture" auditory verbal hallucinations when they are naturally occurring have reported activation of temporal lobe structures during AVHs (12,13). In healthy control subjects, the generation of inner speech, or the silent articulation of words, is associated with activation in the left inferior frontal cortex/insula, the supplementary motor area, the left superior temporal/inferior parietal region, and the right posterior cerebellar region (14). The left inferior frontal cortex has also been shown to be activated during AVHs (12,15). A more recent study suggests the right inferior frontal gyrus may also be involved (16), indicating involvement of inner speech in AVHs. An early positron emission tomography (PET) study investigating auditory verbal imagery—imaging another person's speech—revealed that schizophrenia patients who were prone to auditory hallucinations displayed normal activation of the left inferior frontal gyrus but abnormal activation of the left temporal cortex compared with both schizophrenia patients without a history of AVHs and healthy control subjects (17). More recently, we have used functional magnetic resonance imaging (fMRI) to show normal activation of the left inferior frontal gyrus and attenuation of right temporal cortex activation in hallucination-prone schizophrenia patients, relative to an inner speech task, during verbal imagery (18). To summarize, previous studies indicate that when generating verbal material, schizophrenia patients display normal activation of the inferior frontal cortex but abnormal activity in the temporal cortex, with the changes varying with the condition used as the baseline. However, to demonstrate the specificity of this hypothesis, ideally one would compare simple self-generated speech with comparable external speech to exclude any nonspecific

From the Department of Psychiatry and Neuropsychology (CJPS, M-de-GD, LK), Maastricht University, European Graduate School of Neuroscience, Maastricht; and GGz Eindhoven (CJPS), Eindhoven, The Netherlands; Cognition Schizophrenia and Imaging Laboratory (DKT, KKS, OO, JG, SSS), Division of Psychological Medicine; and Department of Clinical Neuroscience (OO), Institute of Psychiatry, London; and University of Leicester School of Medicine (KKS) Leicester, United Kingdom; and Center Brain and Learning (LK), Department of Psychology and Education, VU University Amsterdam, Amsterdam, The Netherlands.

Address correspondence to Claudia J.P. Simons, M.Sc., Department of Psychiatry and Neuropsychology, Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands; E-mail: c.simons@sp.unimaas.nl.

Received Jun 3, 2009; revised Sep 2, 2009; accepted Sep 3, 2009.

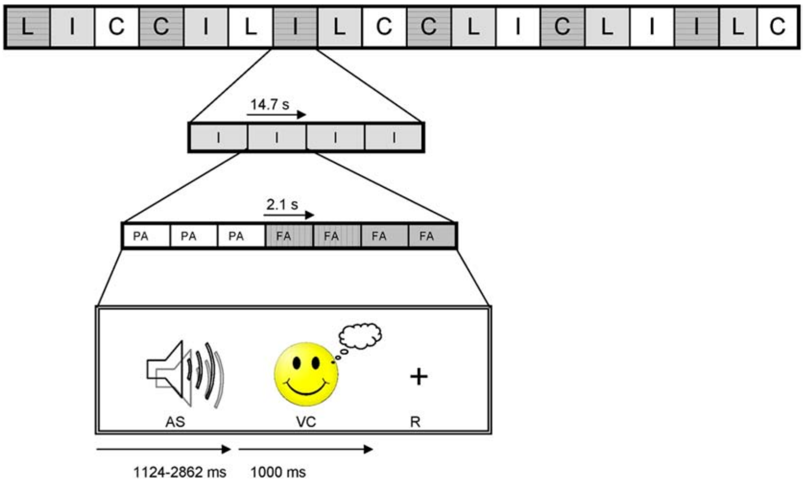

Figure 1. Stimulation protocol. Illustrated are the experimental blocks, the trials within a single block, and how stimulus acquisition, stimulus presentation, and response take place during a trial. AS, auditory stimulus; C, baseline; FA, full acquisition; I, inner speech; L, listening; PA, partial acquisition; R, response; VC, visual cue.

impairments in the auditory cortical processing and to exclude the increased demands associated with generating more complex stimuli such as imagery. To our knowledge, there are no fMRI studies explicitly comparing this in patients with schizophrenia.

In the present study, we used fMRI to investigate brain activation during the perception of internally generated speech and external speech in both patients with schizophrenia and matched healthy control subjects. Activation during the self-monitoring of inner speech was compared directly with activation during listening to externally generated speech, a task requiring no self-monitoring. We predicted that inner speech would lead to attenuation of temporal cortex activation in healthy control subjects as a consequence of corollary discharge, while the patients with schizophrenia would show less attenuated activation of the temporal cortex during inner speech trials, indicating faulty verbal self-monitoring. Since external sensory stimulation does not lead to an efference copy, we anticipated an interaction between task (listening vs. inner speech) and group (control subjects vs. patients) in the temporal cortex. As silent articulation is associated with activation of the left inferior frontal cortex, we also predicted that inner speech trials would be accompanied by an increased activation of this area compared with listening trials in both patients and the control subjects. The present study minimized the effects of acoustic scanner noise during the stimulus presentation through the use of a partially silent acquisition in which the auditory stimulus preceding the inner speech and listening prompts was presented in a partially silent gap followed by a period of continuous image acquisition (19).

Methods and Materials

Subjects

Fifteen male patients with a DSM-IV diagnosis of schizophrenia and right-handed, as assessed with the Annett Handedness Inventory (20), participated in the study. Patients were recruited through consultant and key worker recommendations and had all experienced prominent auditory hallucinations during exacerbations of their illness. Fourteen subjects were outpatients and one subject was an inpatient. All patients were receiving regular doses of antipsychotic medication. Mean age of the patients was 34.7 years (SD = 8.7). Mean duration of illness was 11.2 years (range 3–27). Mean score on the Positive and Negative Syndrome Scale (21) was 48.5 (SD = 16.5, range 30–83). Patients were matched for age, sex, and handedness to a control group. Twelve healthy right-handed, male control subjects were recruited through an advertisement in a city-wide newspaper. They did not suffer from psychiatric disorders and had no family history of psychiatric disorder. Mean age of the control subjects was 34.4 years (SD = 7.9). English was a first language for all subjects and all subjects had a minimum of 11 years of education. Exclusion criteria were any illicit drug use within the previous 6 months or any contraindications to magnetic resonance imaging (MRI) scanning. Potential subjects were assessed on their ability to perform the tasks (detailed below) outside the scanner, approximately 1 week before scanning. Subjects provided written informed consent, and ethical approval was provided by the Institute of Psychiatry and Maudsley National Health Services Trust.

Tasks Performed During fMRI

Subjects performed two active tasks, listening and inner speech, and there was one additional null baseline condition. Each of these tasks was administered over six counterbalanced blocks, each block comprising four listening trials, four inner speech trials, and four baseline trials, with the baseline trials consisting of a silent period equal in length to four paired stimuli (Figure 1).

Auditory stimuli used for the listening and inner speech conditions were 24 neutral sentences, spoken by an adult female native English speaker. During the listening trials, the auditory stimulus was followed by a visual cue prompting the subjects to listen to a second auditory stimulus identical to the first; during the inner speech trials, the auditory stimulus was followed by a visual cue prompting the subjects to covertly imagine repeating the sentence to themselves in their own voice and press a button with their right index finger once this was completed.

Each auditory stimulus was presented once during a listening trial and once during an inner speech trial. Sentences were presented via pneumatically driven earphones, incorporated within ear defenders, specifically designed for functional MRI (Quiet Muff 29 Earmuffs, Avotec, Jensen Beach, Florida). These reduced unattenuated noise from the scanner. As some sentences were longer than others, the duration of the stimuli varied from 1124 msec to 2862 msec, with an average length of 1971 msec. The auditory stimulus was followed by a 1000-msec period in which the visual prompt was presented, followed by the actual task (listening/inner speech). There was a gap before the onset of the next trial (intertrial interval: 14,728 msec). Image acquisi-