Language Processing and Human Voice Perception in Schizophrenia: A Functional Magnetic Resonance Imaging Study

Michihiko Koeda, Hidehiko Takahashi, Noriaki Yahata, Masato Matsuura, Kunihiko Asai, Yoshiro Okubo, and Hiroshi Tanaka

Background: *Neuroimaging studies have demonstrated either reduced left-lateralized activation or reversed language dominance in schizophrenia. These findings of left hemispheric dysfunction could be attributed to language processing tasks, which activate mainly left hemispheric function. Recent functional magnetic resonance imaging studies reported right-lateralized temporal activation by human voice perception, but few studies have investigated activation by human voice in schizophrenia. We aimed to clarify the cerebral function of language processing in schizophrenia patients by considering cerebral activation of human voice perception.*

Methods: *Fourteen right-handed schizophrenia patients and 14 right-handed controls with matched handedness, sex, and education level were scanned by functional magnetic resonance imaging while listening to sentences (SEN), reverse sentences (rSEN), and identifiable non-vocal sounds (SND).*

Results: *Under the SEN-SND and SEN-rSEN contrasts including language processing, patients showed less activation of the left hemisphere than controls in the language-related fronto-tempo-parietal region, hippocampus, thalamus and cingulate gyrus. Under the rSEN-SND contrast including human voice perception, patients showed less activation than controls in the right-lateralized temporal cortices and bilateral posterior cingulate.*

Conclusions: *Our results indicate that schizophrenia patients have impairment of broader bilateral cortical-subcortical regions related to both the semantic network in the left hemisphere and the voice-specific network in the right hemisphere.*

Key Words: Language processing, human voice perception, schizophrenia, semantic network, voice-specific network, functional magnetic resonance imaging (fMRI)

Auditory hallucination and thought disturbance are the main symptoms of schizophrenia and are assumed to be associated with disturbance in language processing (Kircher et al 2002; Mitchell et al 2001; Sommer et al 2001; Woodruff et al 1997). In order to understand the neural basis of these symptoms, achieving the clarification of cerebral function when patients with schizophrenia listen to language will be a challenging target.

Functional magnetic resonance imaging (fMRI) studies have investigated the neural basis of disturbed language processing in schizophrenia (Kasai et al 2003; Kubicki et al 2003; Ragland et al 2004). Functional MRI studies have reported that normal right-handed subjects show left-lateralized activation for language processing (Gaillard et al 2002; Lehericy et al 2000; Schlosser et al 1998), whereas normal left or ambidextrous-handed subjects show bilateral activation for language processing (Hund-Georgiadis et al 2002; Pujol et al 1999; Szaflarski et al 2002). Previous fMRI studies with language listening tasks have demonstrated that schizophrenia patients show either reduced left hemispheric activation (Kiehl and Liddle 2001; Kircher et al 2001) or reversed language dominance (Menon et al 2001; Ngan et al 2003; Woodruff et al 1997), and schizophrenia patients have disturbed left hemisphere dominance for language processing (Sommer et al 2001, 2003). Similarly, fMRI studies with verbal fluency tasks have shown either reduced left frontal activation (Artiges et al 2000; Curtis et al 1998; Yurgelun-Todd et al 1996) or reversed activation in the frontal cortices (Crow 2000; Sommer et al 2001, 2003). These findings of hypo or reversed activation of schizophrenia could depend on the nature of language processing tasks, which mainly activate the left hemisphere in normal subjects.

In schizophrenia, investigating cerebral response to human voice is significant since human voice perception is known to closely associate with the generative mechanism of functional auditory hallucination like languages or specific identifiable sounds (Hayward 2003; Hunter and Woodruff 2004). Recent fMRI studies have shown that the human-voice-specific area is located in the superior temporal sulcus (STS) with right hemisphere dominance in normal subjects (Belin and Zatorre 2003; Belin et al 2002; Belin et al 2000). When auditory hallucination appears, patients with schizophrenia have demonstrated increasing cerebral activation in the temporal cortex (Bentaleb et al 2002; Dierks et al 1999; Woodruff et al 1997). If cerebral activation in the right hemisphere is increased by human voice perception, language dominance in the temporal cortices could be reversed when schizophrenia patients are listening to language. Recent studies have indicated the importance of investigating right-hemisphere language function as a social perception in schizophrenia (Abdi and Sharma 2004; Mitchell and Crow 2005; Onitsuka et al 2005; Williams et al 2004). However, to our knowledge, no studies concerning language processing of schizophrenia have taken cerebral activation by human voice perception into account. Therefore, we investigated how cerebral activation for language processing and human voice perception in schizophrenia patients are different in comparison to normal subjects.

From the Department of Bioinformatics (MK, HTan), Medical Research Institute; Department of Bioinformatics and Computational Biology (HTan), School of Biomedical Sciences; Department of Biofunctional Informatics (MM), Tokyo Medical and Dental University; Department of Pharmacology (NY) and the Department of Neuropsychiatry (YO), Nippon Medical School, Tokyo; National Institute of Radiological Sciences Brain Imaging Project (HTak), Chiba; Asai Hospital (KA), Togane, Japan.

Address reprint requests to Yoshiro Okubo, M.D., Ph.D., Department of Neuropsychiatry, Nippon Medical School, 1-1-5, Sendagi, Bunkyo-ku, Tokyo, 113-8603 Japan; E-mail: okubo-y@nms.ac.jp.

Received August 11, 2005; revised January 6, 2006; accepted January 13, 2006.

0006-3223/06/$32.00
doi:10.1016/j.biopsych.2006.01.013

The aim of our research was to clarify cerebral function by language processing in patients with schizophrenia by considering activation by human voice perception.

Methods and Materials

Subjects

Fourteen schizophrenia patients (12 males and 2 females, mean age 31.6 years, SD = 7.0) meeting the DSM-IV criteria for schizophrenia were studied. Diagnoses were made by HTak, YO, and the attending psychiatrists on the basis of a review of their charts and a conventionally semi-structured interview. Exclusion criteria were current or past substance abuse and a history of alcohol-related problems, mood disorder, or organic brain disease. Eleven patients were recruited from the outpatient unit of Asai Hospital, and 3 were recruited from the inpatient unit. Ten of the 14 patients were the same as in our previous fMRI studies investigating emotional neural responses (Takahashi et al 2004). As for the subtypes of schizophrenia, 13 patients were paranoid and one had undifferentiated schizophrenia. Thirteen of the 14 patients were receiving neuroleptics (mean risperidone equivalent daily dosage = 3.6 mg, SD = 3.5; 8 patients, risperidone; 1 patient, perospirone; 1 patient, olanzapine; 1 patient, zotepine; 2 patients, sulpiride), and one was not receiving any neuroleptics. Mean illness duration was 9.6 (SD = 9.7) years. Clinical symptoms were assessed by the Brief Psychiatric Rating Scale (BPRS) (Overall and Gorham 1962). The ratings were reviewed by HTak and YO after the patient interview, and disagreements were resolved by consensus; consensus ratings were used in this study. Further, sum scores for positive and negative symptoms were calculated, with the positive symptom subscale including the following eight items: conceptual disorganization, mannerisms and posturing, hostility, grandiosity, suspiciousness, hallucinatory behavior, unusual thought content, and excitement. The negative symptom subscale included these three items: emotional withdrawal, motor retardation, and blunted affect. The mean score of BPRS was 13.9 (SD = 7.4). The mean positive symptom score was 1.9 (SD = 2.4), and the mean negative symptom score was 3.6 (SD = 2.9). The control group consisted of 14 normal subjects (10 males and 4 females, mean age 29.1 years, SD = 7.8), who were recruited from the surrounding community. The candidates were carefully screened and standardized interviews were conducted by trained psychiatrists (HTak and YO). They did not meet the criteria for any psychiatric disorders. None of the control subjects was taking alcohol or medication at the time, nor did they have a history of psychiatric disorder, significant physical illness, head injury, neurological disorder, or alcohol or drug dependence. Patients tended to have a lower educational status but there was no significant difference in the mean period of education between the controls and patients (mean ± SD; patients 13.2 ± 2.0 years, control subjects 14.4 ± 1.8 years; $p = .15$, *t*-test). All of the patients and control subjects were right-handed, as investigated by the Edinburgh Handedness Inventory (EHI) (mean ± SD; control subjects 92.3 ± 10.3; patients 91.5 ± 9.9, $p > .05$, *t*-test) (Oldfield 1971). Previous fMRI studies defined a right-handed subject as a subject with an EHI score of more than 50 or 52 (Springer et al 1999; Szaflarski et al 2002). Therefore, we defined a right-handed subject as "a subject with an EHI score of more than 52." In our study, 16 patients were initially included, but two ambidextrous patients (EHI: less than 52) were then excluded from the analysis. There was no significant difference in the mean of the EHI score between the two groups. They all underwent MRI to rule out cerebral anatomic abnormalities. After the procedures had been fully explained to the subjects, written informed consent was obtained in accordance with the guidelines of the Asai Hospital Ethics Committee.

Experimental Design

In a single session, three types of stimuli were presented: forward-played sentences (SEN); reverse sentences (rSEN), the same sentences, but played in reverse; and identifiable nonvocal sounds (SND). The duration of each stimulus was 20 sec, and rSEN, SND, and SEN were played in sequence to each subject. Before each sound, the subjects listened to silence from the headphones for 20 sec (rest condition). Each set was 120 sec, and consisted of these three sound conditions and the rest condition. One session consisted of 8 sets, with a total scanning time of 960 sec (Figure 1). As identifiable non-vocal sounds for the SND condition, sounds of a shower, washing machine, bell, and computer printer were used. The sentences were in Japanese, spoken by two speakers, one male and one female, and represented a single topic per set, with one session consisting of four topics, each repeated twice randomly. Concerning the contents of the sentences, each topic was expressed by one or two sentences, consisting of 6–7 phrases including compound sentences. These sentences used conjunctional phrases or long adjuncts. Therefore, each subject was required to comprehend complex situations and understand the connection of the phrases or sentences (Appendix 1). In our experiment, two voices, one male and one female, were used alternately for the sound contents of sentences and reverse sentences. The material of the sentences included the linguistic section of the contents of Wechsler's Memory Scale – Revised, translated into Japanese (Sugishita 2001; Wechsler 1987). In our research, we used Japanese sentences for Japanese native subjects because research of semantic processing of language has shown left-

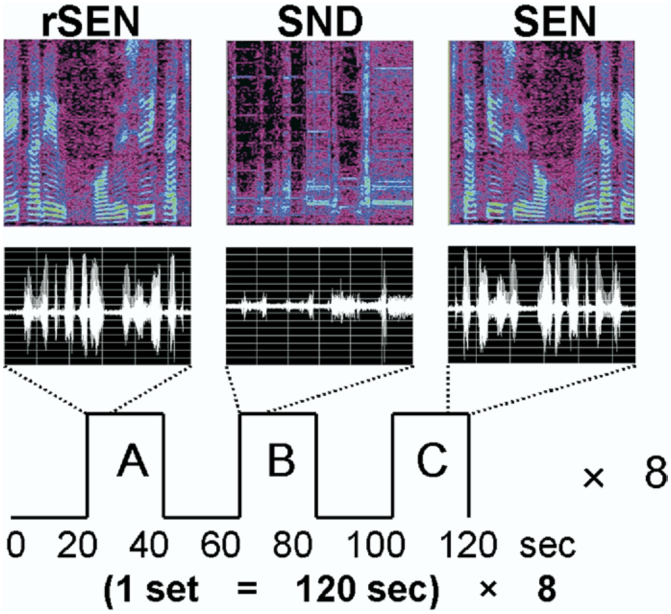

Figure 1. (A) Reverse sentences (rSEN), **(B)** Identifiable non-vocal sounds (SND), **(C)** Sentences (SEN). The top row shows sound spectrograms under the three sound conditions. The horizontal axis shows the time domain. The vertical axis shows the frequency (about 15–20000 Hz) of the tone domain. The middle row shows the time-domain waveforms. The horizontal axis shows the tone domain. The vertical axis shows the power of the sound.