Voice Identity Recognition Failure in Patients With Schizophrenia

Lucy Alba-Ferrara, PhD,† Susanne Weis, PhD,† Ljubica Damjanovic, PhD,‡ Matthew Rowett, MD,§ and Markus Hausmann, PhD†*

Abstract: Cognitive models propose that auditory verbal hallucinations arise through inner speech misidentification. However, such models cannot explain why the voices in hallucinations often have identities different from the hearer. This study investigated whether a general voice identity recognition difficulty might be present in schizophrenia and related to auditory verbal hallucinations. Twenty-five schizophrenia patients and 13 healthy controls were tested on recognition of famous voices. Signal detection theory was used to calculate perceptual sensitivity and response criterion measures. Schizophrenia patients obtained fewer hits and had lower perceptual sensitivity to detect famous voices than healthy controls did. There were no differences between groups in false alarm rate or response criterion. A symptom-based analysis demonstrated that especially those patients with auditory verbal hallucinations performed poorly in the task. The results indicate that patients with hallucinations are impaired at voice identity recognition because of decreased sensitivity, which may result in inner speech misidentification.

Key Words: Phonagnosia, externalizing bias, familiarity, delusion, hallucination.

(*J Nerv Ment Dis* 2012;200: 784–790)

Auditory verbal hallucinations (AVHs) are one of the most striking symptoms of schizophrenia, affecting up to 70% of schizophrenia patients during the course of the illness (Bentall, 1990). Despite the vast amount of research carried out in the area, the mechanisms of formation of AVH are still poorly understood. A prominent phenomenological feature of AVH is the perception of voices that have a specific identity (Stephane et al., 2003). Moreover, the voices seem to be generated by a person other than the self and often have the acoustical features (*e.g.*, pitch, tempo, amplitude, and even accent) of a particular individual, different to the hearer's own (Jones and Fernyhough, 2007). The voice identity specificity of AVH challenged theories claiming a misattribution of inner speech as the foundations of AVH, as these theories still need to account for the mechanism by which inner speech conveys acoustic properties different from the hearer's own voice. The missing link could be an additional voice identity recognition deficit in patients with AVH, as difficulties in recognizing one's own voice could contribute to the attribution of self-generated material to an external source (Alba-Ferrara et al., 2012).

Previous research has intended to address whether AVHs result from impaired self-monitoring of verbal material (Johns et al., 2001). In the study by Johns et al. (2001), AVH patients and controls were asked to read words aloud while wearing headphones that transmitted the vocal input back to the participant. In some of the trials, the transmission of the speech was distorted (acoustic features such as pitch were modified). In other trials, participants heard someone else's voice instead of their own as they spoke. Control trials were also included in which participants could hear their voice without any modification. Immediately after saying the words, participants identified the source of the voice they heard as own, alien, or unsure via a key press. The results showed that patients with positive symptoms, compared with patients without positive symptoms and healthy controls, misidentify their own voice as that of someone else when presented during the distorted feedback trials (Johns et al., 2001). This finding was interpreted by the authors in terms of abnormal self-monitoring. However, there might be alternative explanations of this misidentification of self-generated speech.

For example, Allen et al. (2004) tested AVH patients with distorted recordings of adjectives spoken in their own or other person's voice. Because this paradigm did not involve generating verbal material by the subjects at the moment of testing, the task could be performed without the use of verbal self-monitoring. The authors propose that the misattribution was due to an externalizing bias when processing unusual perceptual information. Finally, the authors dismissed the idea that the misattribution was due to a general deficit of voice discrimination, arguing that AVH patients tend to attribute the utterances to external sources instead of choosing the options "their own voice" or "unsure." However, it is important to consider that one's voice sounds different when it is heard from a recording (Békésy, 1949). During self-generated speech, sound reaches the inner ear by way of two separate paths. Air-conducted sound is transmitted from the surrounding environment through the external auditory canal, eardrum, and middle ear to the cochlea. Bone-conducted sound reaches the cochlea directly through the tissues of the head. The voice heard during self-generated speech is perceived by the combination of sound carried along both paths, resulting in a deeper and more resonant sound. Instead, listening to an external sound is performed through the air-conducted path solely. Thus, in the study of Allen et al., the misattribution bias might be explained by taking into account that the participant's own voice reproduced by an external device sounds different in comparison with self-speech production. In fact, the loss of spectral information caused by the applied pitch distortion results in stimuli that are harder to recognize. Moreover, the recorded voice is not transmitted by bone conduction, resulting in a mismatch between the internal representation of one's own voice and perception of the recorded voice. The mentioned loss of acoustical information of the recorded stimuli, in addition to the voice identity recognition deficit in AVH, would result in the misattribution of the stimuli as alien. Although controls might be able to compensate for the mismatch of their voices with the recorded stimuli, the task might become critical for those with a voice discrimination deficit. Thus, applying a voice identity recognition paradigm in AVH patients free from the bone-conducted pathway confound would show whether there is a genuine voice misidentification deficit in this population.

A recent study has assessed voice recognition using a novel paradigm (Zhang et al., 2008). This study presented AVH patients with personally familiar voices and with voices of strangers. Participants had to decide whether the voices were familiar or unfamiliar. This study found impairment in voice recognition in AVH patients in comparison with non-AVH patients and healthy controls. Unfortunately, the authors did not test whether the patients could recognize

*Roskamp Laboratory of Brain Development, Modulation and Repair, Department of Psychiatry and Neurosciences, Morsani College of Medicine, University of South Florida; †Department of Psychology, Durham University, Durham; ‡Department of Psychology, University of Chester, Chester; and §Walkergate Park Hospital, Newcastle-Upon-Tyne, UK.

Send reprint requests to Lucy Alba-Ferrara, PhD, Roskamp Laboratory of Brain Development, Modulation and Repair, Department of Psychiatry and Neurosciences, Morsani College of Medicine, University of South Florida, 3515 E Fletcher Avenue, Tampa, FL 33613. E-mail: lalbafer@health.usf.edu.

ISSN: 0022-3018/12/20009–0784
DOI: 10.1097/NMD.0b013e318266f835

the identity of the speaker, instead they assessed only familiarity. Thus, it is not entirely clear whether a voice identity recognition difficulty (phonagnosia) is present in this group of patients or if, alternatively, the results indicate a voice familiarity problem. In addition, because signal detection measures were not calculated, it is difficult to infer from Zhang et al. (2008) whether the differences in performance between AVH patients and controls are due to a decreased sensitivity to detect familiar over unfamiliar voices or to a bias toward a conservative approach at solving the task, such as a tendency to classify any voice as unfamiliar (Snodgrass and Corwin, 1988). If AVH patients cannot distinguish the identity of the speaker by the tone of voice, it may be the case that even their own voice may not be recognized, resulting in the source of the voice being perceived as alien.

A different strand of research proposes that patients' inability to recognize their own voices may be due to a general self-agnosia, that is, impairment at distinguishing between self and others (Waters and Badcock, 2010). This idea has been investigated with some memory paradigms. One of the studies in this topic assessed source-based memory with a task in which schizophrenia patients were asked to identify whether a word was previously heard in a female or male voice (Weiss et al., 2008). The authors did not find differences between patients and controls and interpreted the negative finding as related to the fact that the assessed patients had only low levels of psychopathology with predominantly negative symptoms. Another study demonstrated that schizophrenia patients have difficulties remembering self-generated thought words. Here, the performance of patients with positive symptoms was particularly poor (Keefe et al., 2002). Finally, a more recent study assessed sex identity recognition in schizophrenia using vocal stimuli (Waters and Badcock, 2009). This study found that, compared with controls, patients were impaired at memory only for female but not for male voices. Surprisingly, this study did not find differences between patients with more pronounced auditory hallucinations in comparison with patients with low psychopathology.

It is the aim of the present study to establish whether schizophrenia patients and particularly those who have AVH are impaired in voice identity recognition by using an established paradigm in the assessment of voice identity recall as well as of phonagnosia (Damjanovic and Hanley, 2007; Garrido et al., 2009; Hanley and Damjanovic, 2009). The assessment of voice recognition in schizophrenia might be particularly important, although self-recognition deficits in these patients occur in all modalities (Waters et al., 2010). The hallucinations that patients experience are predominantly auditory, suggesting that the vocal channel is particularly compromised. Therefore, it is hypothesized that schizophrenia patients, and particularly the subgroup with prominent positive symptoms, will identify fewer famous voices and will find the voices unfamiliar, independently of whether they belonged to known or unknown speakers, in comparison with non-AVH patients and healthy controls.

METHODS

Participants

Twenty-five (20 men) individuals who met the *DSM-IV-TR* criteria (American Psychiatric Association, 2000) for schizophrenia were recruited from several outpatient clinics from Northumberland, Tyne, and Wear National Health System (NHS) Foundation Trust and Tees, Esk, and Wear Valleys NHS Foundation Trust. The psychiatric diagnosis was confirmed by an independent psychiatrist. All patients were taking antipsychotic drugs such as haloperidol (n = 1), flupentixol (n = 5), risperidone (n = 7), olanzapine (n = 3), aripiprazole (n = 4), or clozapine (n = 4) (for a comparison of medication dosage between patients, see Table 2). Exclusion criteria for patients were multiple diagnoses such as the presence of comorbidities with axis I disorders of the *DSM* or existence of a neurological condition.

In addition, 13 healthy participants (8 men) were recruited via advertisement in the local post office. The leaflet for the recruitment asked for participants to take part in psychology experiments in exchange for a £30 compensation reward. To avoid a potential self-selection bias, participants were informed about the nature of the experiment after recruitment. Participants were screened for history of psychiatric illness, head injury, tinnitus, epilepsy, and drug use. All patients and controls were native English speakers and were permanent residents in the UK. After receiving a detailed description of the study, written informed consent was obtained from each participant. The study was approved by the regional NHS ethics committee and Durham University Ethics Advisory Committee. Subjects received £30 for participating in the study.

At the beginning of the testing session, all participants completed a hearing screening and were assessed with the National Adult Reading Test (NART; Nelson and Willson, 1991), a test devised to estimate premorbid intellectual performance (IQ) with a high test-retest reliability (Morrison et al., 2000). All participants reported that they did not have hearing impairments or tinnitus. An additional hearing test was conducted using monaural white-noise bursts (duration, 1 second), presented via headphones with various sound-pressure levels (steps of 10 dB). More details about the hearing test can be found in Hirnstein et al. (2007). This test revealed normal performance in all participants. Moreover, there were no differences in handedness and verbal IQ between groups. The results of the NART, hearing test, and the Edinburgh handedness questionnaire are shown in Table 1.

Assessment of Psychopathology

Interviews were conducted by a qualified clinical psychologist using a semistructured diagnostic interview, the Comprehensive Assessment of Symptoms and History (Andreasen et al., 1992). This interview includes the Scale for the Assessment of Positive Symptoms (SAPS; with 34 items measured on an ordinal scale ranging from 0 [absent] to 5 [severe]) (Andreasen, 1984b) and the Scale for the Assessment of Negative Symptoms (with 21 items) (Andreasen, 1984a). Details about the scales can be found in Table 2.

Twelve patients who were not currently experiencing hallucinations (as defined by a score of ≤1 in the SAPS hallucination global score) were allocated to the non-hallucinators group (NAVH). Patients who reported hallucinations (scoring at least 3 on the SAPS hallucinations global score) were allocated to the hallucinators group (AVH). None of the patients scored between 1 and 3 in this scale. The AVH group subsequently completed the auditory hallucination subscale corresponding to the Psychotic Symptom Rating Scales (Haddock et al., 1999). This subscale consists of 11 items measuring frequency, duration, severity, and intensity of distress caused by auditory hallucinations, as well as the controllability, loudness, location, negative content, degree of negative content, beliefs about origin of voices, and disruption they cause in daily life. A 5-point ordinal scale is used to rate symptom scores (0–4).

The three groups did not differ significantly in education, age, or verbal IQ (see Table 1). However, there were differences between the NAVH and AVH groups in their mean SAPS global scores for delusion as demonstrated using Mann-Whitney U-test (U = 26.5, $p < 0.005$) (Table 2). For the psychopathology variables, a nonparametric test was used because the assumption of normal distribution was violated as demonstrated by the Kolmogorov-Smirnov test for one sample. In addition, after converting the medication dosage into chlorpromazine equivalents (Andreasen et al., 2010), an unpaired t-test revealed significantly higher medication dosage in the AVH group (mean [SD], 457 [121]) in comparison with the NAVH group (mean [SD], 284 [140]; $t_7 = 2.65$, $p < 0.033$).

TABLE 1. Demographic Features and Hearing Test

Demographic Data	Groups, Mean (SD) AVH (n = 13; 10 Men)	Non-AVH (n = 12; 10 Men)	Controls (n = 13; 9 Men)	Analysis (One-Way ANOVA)
Age (yr)	41.73 (2.62)	37.83 (2.87)	42.69 (3.09)	$p = 0.467$
NART/verbal IQ	112.14 (1.79)	110.42 (1.74)	112.69 (1.24)	$p = 0.573$
Hearing test	50.00 (1.24)	55.00 (1.82)	50.00 (1.16)	$p = 0.735$
Years of education	12.73 (2.02)	13.08 (2.39)	14.06 (2.93)	$p = 0.318$

Characteristics of patient groups and controls. The three groups, schizophrenia patients with hallucinations (AVH), schizophrenia patients without hallucinations (non-AVH), and healthy controls, did not differ significantly in education, age, or verbal IQ. Absolute hearing threshold is expressed in dB SPL.

ANOVA indicates analysis of variance; NART, National Adult Reading Test.

Materials

The stimuli used in the experimental task were taken from the Damjanovic Famous Voices data set (Damjanovic and Hanley, 2007). They consisted of 96 voice samples (half female) taken from television interviews, each of which lasted approximately 7 seconds. Half of the samples were from famous people, such as David Beckham and Margaret Thatcher, and the other half were from nonfamous people. The tone of voice of each sample was neutral, and its content did not offer any clues regarding the identity or occupation of the speaker. The audio tracks from both famous and nonfamous voices were presented via headphones. The whole task lasted approximately 25 minutes. Participants could opt to take breaks during the task. Stimuli were presented using E-prime.

Procedure

The participants entered a quiet testing room and were seated in front of a laptop. They were informed that they would be presented with a sequence of voices, some of which belonged to people who were well known in the British media, whereas others belonged to nonfamous individuals. For each voice, participants had to decide by key press whether the voice belonged to somebody famous or not. If they considered the voice to be famous, they were asked to orally classify their responses as one of the following categories: remember (R), know (K), or guess (G). Along the lines of previous research (Maylor, 1995), the experimenter asked the participants if they could recall the person's name (R) or if they could associate some facts about the person (*e.g.*, profession or recall of an event the celebrity has taken part in) even though they could not recall the person's name (K). If participants thought they have heard the voice before but they could not recall anything about the person, this was registered by the experimenter as G. Including G as a possible response is necessary to separate confident from unconfident K responses by giving the subject the possibility to answer G in case of very low confidence. However, G responses should be interpreted with caution because they do not strictly reflect recognition (Gardiner et al., 2002).

Analysis

For the statistical analyses, only those trials where participants expressed confidence in their memory (correct R and K responses) were included. All responses were converted to proportions by dividing the number of responses per category (R and K) by 48 (total of famous voices). The total number of hits was defined as the sum of

TABLE 2. Symptoms

Symptoms Rating	Group, Mean (SD) AVH (n = 13; 10 Men)	Non-AVH (n = 12; 10 Men)	Analysis (Mann-Whitney U-Test)
Duration of illness	13.73 (2.22)	15.17 (2.01)	$U = 86$, $p = 0.844$
SANS total	10.88 (1.51)	10.04 (0.84)	$U = 77$, $p = 0.525$
Affective flattening	1.53 (0.25)	1.71 (0.4)	$U = 86$, $p = 0.862$
Alogia	1.43 (0.27)	1.79 (0.47)	$U = 81$, $p = 0.655$
Avolition	2.27 (0.27)	2.29 (0.44)	$U = 89$, $p = 0.98$
Anhedonia	2.57 (0.25)	2.17 (0.42)	$U = 78$, $p = 0.552$
Attention	2.6 (0.3)	2.92 (0.43)	$U = 70$, $p = 0.330$
SAPS total	9.9 (0.69)	4.92 (0.87)	$U = 21$, $p = 0.001$*
Hallucinations	3.83 (0.24)	0.96 (0.23)	$U = 1.5$, $p < 0.001$**
Delusions	2.97 (0.33)	1.17 (0.31)	$U = 26.5$, $p = 0.002$*
Bizarre behavior	1.47 (0.25)	1.42 (0.31)	$U = 86$, $p = 0.840$
Positive formal thought	1.38 (0.33)	1.63 (0.3)	$U = 79.5$, $p = 0.598$
PSYRATS (hallucination subscale)	25.87 (1.73)	1.75 (1.18)	$U = 2$, $p = <0.001$**
Chlorpromazine equivalent[a]	457 (121)	284 (140)	$t = 2.65$, $p = 0.033$*

[a]Drug doses were converted into clozapine equivalents using the formula in Andreasen et al. (2010).

*Reported values significant at $P < .05$.

**Reported values significant at $P < .001$.

AVHs (n = 13) who were not experiencing hallucinations during the testing session, as defined by a score of 1 or below in the SAPS hallucinations global score. Non-AVHs (n = 13) scored at least 3 in the SAPS hallucinations global score. Between-groups comparisons analyzed with nonparametric Mann-Whitney U-test as the sample were not normally distributed with respect to their symptoms.

AVH indicates schizophrenia patients with hallucinations; non-AVH, schizophrenia patients without hallucinations; SANS, Scale for the Assessment of Negative Symptoms; SAPS, Scale for the Assessment of Positive Symptoms; PSYRATS, Psychotic Symptom Rating Scales.

R and K responses for each participant, and the total number of false alarms (FAs) comprised all nonfamous voices categorized as famous. G responses were excluded from the analyses.

We calculated signal detection theory measures to quantify the participant's ability to discriminate famous from nonfamous voices and their response criterion in answering.

The sensitivity index A′ was calculated as follows (Donaldson, 1996):

$$A' = \frac{1}{2} + \frac{(HIT - FA)(1 + HIT - FA)}{4\,HIT(1 - FA)}.$$

The A′ index, an indicator of discrimination of famous over nonfamous voices, can vary between 0 and 1, with values of 1 indicating perfect discrimination of famous from nonfamous voices and values around 0.5 indicating chance performance. The Aaronson and Watts (1987) modified formula was applied to cases where FA rates exceeded hit. This correction fixes the lower bound at 0.5; thus, in this case, A′ varies between 0.5 and 1.

The response criterion B''_D was calculated as follows:

$$B''_D = \frac{(1 - HIT)(1 - FA) - (HIT)(FA)}{(1 - HIT)(1 - FA) + (HIT)(FA)}.$$

This index provides a measure of the participant's tendency to categorize each voice as famous. The values for B''_D can vary between −1, a response bias of classifying a voice as famous (lax criterion), and +1, a response bias of classifying a voice as nonfamous (strict criterion), with a value of 0 representing a neutral bias.

Hits, FAs, A′ and B''_D were each subjected to a 2 × 2 analysis of variance (ANOVA) with sex of voice (male voices, female voices) as within-subject variable and group (schizophrenia patients, healthy controls) as between-subject variable. For a complementary symptom-based analysis, a 2 × 3 ANOVA was calculated with the only difference that the between-subjects variable had three levels (healthy controls, AVH schizophrenia patients, and NAVH schizophrenia patients).

Finally, an additional medication analysis was performed in patients for the measures in which both patient groups differed. This analysis comprised analysis of covariance (ANCOVA) with the

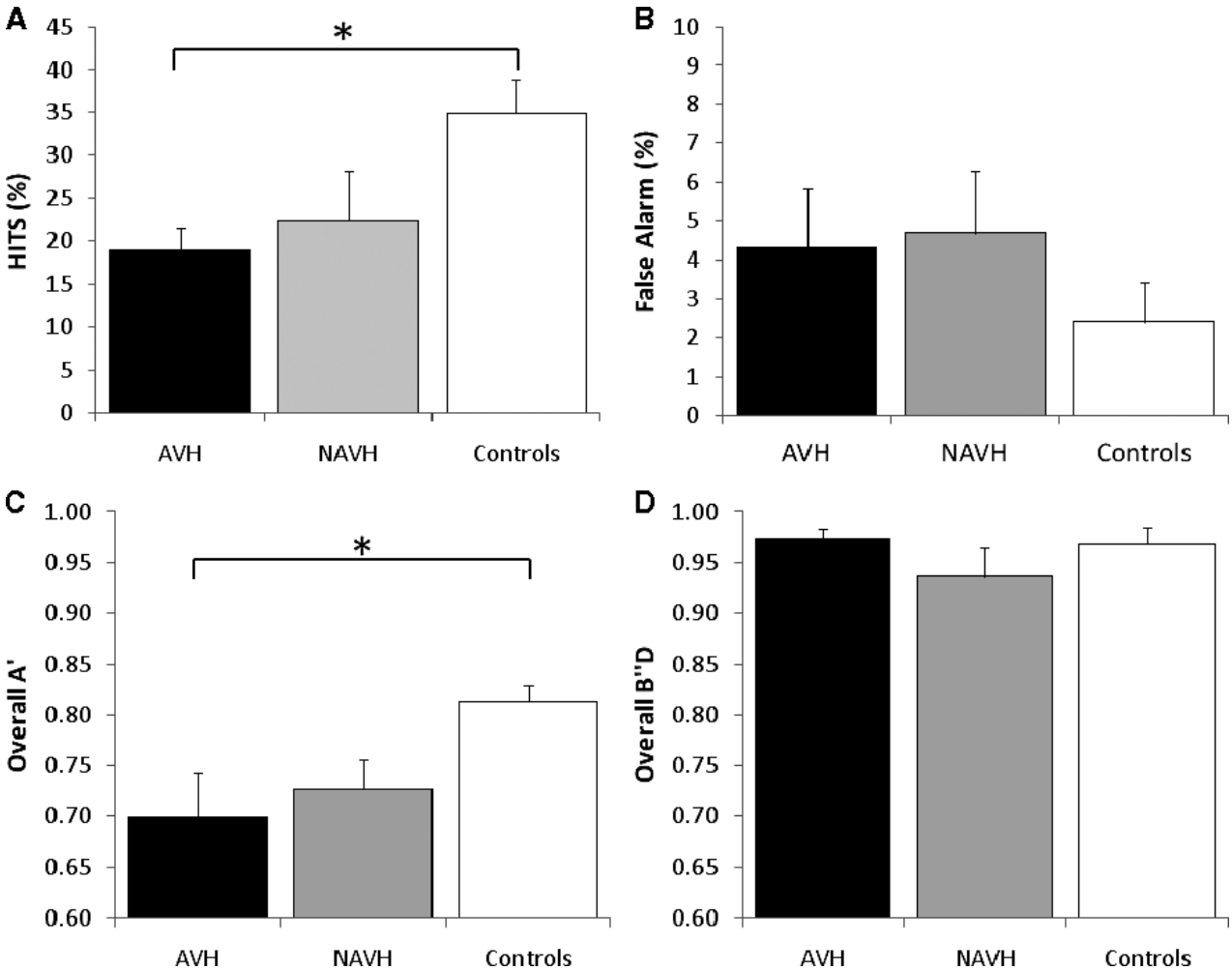

FIGURE 1. Performance and signal detection theory measures for the schizophrenia patients with hallucinations (AVH), healthy controls (controls), and schizophrenia patients without hallucinations (NAVH). Error bars depict standard errors, *$p < 0.005$. (A) Mean hits (famous voices correctly identified). (B) Mean false alarms (nonfamous voices incorrectly identified). (C) Mean discrimination index A′. (D) Mean response criterion B''_D

measure in which patient groups differed as a dependent variable, with groups as a between-subjects variable (AVH, NAVH) and with medication dosage as a covariate. This analysis was sought to extricate the effect of medication from the task performance.

RESULTS

All groups were carefully matched for age and IQ. If we additionally include age and IQ as potential confounders/covariates, none of the findings reported here changes significantly.

Hit Rates

The 2 × 2 ANOVA with group as between-subject factor and sex of voice as within-subject factor revealed a significant main effect of sex of voice ($F_{1,38} = 25.71$; $p = 0.001$), indicating more hits for male voices (mean, 33.53; SE, 3.27) than for female voices (mean, 21.13; SE, 2.03). There was no interaction between sex of voice and group ($F_{1,38} = 0.01$; $p = 0.948$). The main effect of group was significant ($F_{1,33} = 5.13$; $p = 0.003$), indicating more hits in the control group (mean, 32.53; SE, 4.35) than in the patient group (mean, 20.38; SE, 2.97). To disentangle whether this main effect was associated to the symptoms of the patients, an additional 2 × 3 ANOVA was performed in which group was divided into three subgroups (AVH, NAVH, and controls). The main effect of group was again significant ($F_{2,35} = 3.96$, $p < 0.018$). Alpha-corrected post hoc tests (Sidak) showed that the AVH patient group obtained less hits (mean, 18.9; SE, 2.5; $p < 0.012$) than healthy controls did (mean, 34.94; SE, 4.25). The NAVH group (mean, 22.4; SE, 5.91) did not differ from the AVH group ($t_{23} = -0.56$, $p = 0.581$) and from controls ($t_{23} = 1.79$, $p = 0.087$; Figure 1). Finally, to assess whether the differences between patient groups was due to the differences in medication dosage, a 2 × 2 ANCOVA was performed on patients only and dosage of medication was included as potential confounder. The results did not reveal any effect or interaction of medication dosage (all $F_{1,16} < 3.52$; $p = 0.080$).

False Alarms

The 2 × 2 ANOVA did not reveal any significant main effects or interaction (all $F \leq 2.84$, all $p \geq 0.110$). To test if there was any effect associated with symptoms, an additional 2 × 3 ANOVA was performed in which group was divided in three subgroups (AVH, NAVH, and controls). This analysis also did not reveal a main effect nor interaction with group.

Sensitivity Index A′

The ANOVA revealed a significant main effect of sex of voice ($F_{1,35} = 8.32$; $p = 0.007$), indicating higher sensitivity for male voices (mean, 0.81; SE, 0.02) than for female voices (mean, 0.76; SE, 0.01). There was no interaction between sex of voice and group ($F_{1,35} = 0.39$; $p = 0.538$). However, there was a significant main effect of group ($F_{1,35} = 9.33$; $p = 0.004$), indicating lower sensitivity in patients (mean, 0.75; SE, 0.014) than in the control group (mean, 0.82; SE, 0.02). Again, to disentangle whether this main effect was also associated with the symptoms of the patients, a 2 × 3 ANOVA was performed with the group divided into three subgroups (AVH, NAVH, and controls). There was a significant main effect of group ($F_{1,34} = 4.53$; $p = 0.018$). Post hoc Sidak corrected comparisons showed a significant difference between the AVH group (mean, 0.69; SE, 0.02) and healthy controls (mean, 0.81; SE, 0.01; $p = 0.035$). The non-AVH group (mean, 0.73; SE, 0.02) did not differ from the AVH group ($t_{23} = -0.51$, $p = 0.611$) and healthy controls ($t_{23} = 2.60$, $p = 0.198$). Finally, to assess whether the differences between patient groups were confounded by medication dosage, an ANCOVA was performed on patients only and dosage of medication was included as potential confounder. The results did not find any significant effects of the covariate (all $F_{1,14} \leq 0.01$; $p \geq 0.957$).

Response Criterion B''_D

For B''_D, the 2 × 2 ANOVA did not reveal any significant effects (all $F \leq 1.52$, all $p \geq 0.226$). Only the main effect of group approached significance ($F_{1,30} = 3.78$; $p = 0.061$), indicating a trend of a stronger response bias in patients toward a more conservative criterion (mean, 0.77; SE, 0.03) than in healthy controls (mean, 0.66; SE, 0.04). To test if there was any effect associated with symptoms, an additional 2 × 3 ANOVA was performed in which group was divided into three subgroups (AVH, NAVH, and controls). Again, only the main effect of group approached significance ($F_{2,34} = 3.19$; $p = 0.061$), and post hoc comparisons showed only a trend by which the AVH group (mean, 0.80; SE, 0.06) might have adopted a more conservative criterion in comparison with the control group (mean, 0.66; SE, 0.040; $p = 0.058$). However, this difference was not significant.

DISCUSSION

In summary, this study found significant differences between groups in the percentage of hits obtained in the famous voices recognition task, as well as in the sensitivity measure A′. That is, schizophrenia patients obtained a lower percentage of hits and had less sensitivity to identify famous voices than healthy controls did. When splitting the patient group between AVH and NAVH patients, the analysis revealed that AVH patients obtained significantly lower hit rates and sensitivity index A′ than healthy controls did. AVH patients also showed a trend toward obtaining lower hit rates and A′ than NAVH patients. The healthy control group did not significantly differ from the NAVH group. On the contrary, there were no differences between schizophrenia patients and healthy controls in percentage of FAs, neither in the response criterion B''_D measure.

The main finding from the present study shows significant differences between groups in voice identity recognition. Because the results of the healthy control group replicate those of previous studies using the same stimuli in terms of hit rates and sensitivity index A′ for famous voices (Damjanovic and Hanley, 2007; Garrido et al., 2009), schizophrenia patients were found to be impaired in this capacity. A possible interpretation for these findings is that schizophrenia patients are less accurate than healthy controls at extracting a signal (a particular famous voice) from a noisy environment (small individual differences of acoustical features in voices). Thus, this interpretation implies that the patient group is characterized by low signal-to-noise internal representations of individual voices. It has been demonstrated that in schizophrenia, there are high levels of internal noise, which results in reduced signal-to-noise ratio, leading to poor prosodic stimuli recognition (Bach et al., 2008). Both emotional prosody and voice identity recognition rely, in part, on pitch analysis (Perrot et al., 2007). In fact, difficulties in low-level auditory perception such as pitch discrimination have also been found in schizophrenia (Leitman et al., 2005, 2007; Matsumoto et al., 2006). Moreover, pitch perception difficulties in schizophrenia patients have been shown to impact negatively on higher auditory processing such as prosody perception (Leitman et al., 2011). Thus, it might be the case that voice perception difficulties in schizophrenia patients are rooted in pitch perception difficulties.

It should be noted that, although the whole patient sample showed difficulties compared with controls at recognizing famous voices, the subgroup of patients with AVH seems to perform particularly poorly in comparison with controls. Such a finding is in opposition to that of Waters and Badcock (2009), who did not find AVH patients to be more impaired than NAVH patients. To divide groups, Waters and Badcock applied a cutoff score of 2 or more in the SAPS

hallucination subscale. Instead, the present study used a cutoff score of 3 or more to allocate patients to the AVH subgroup. Thus, patients of AVH group in the present study were probably more extreme in their symptoms and thus more likely to show the voice identity recognition impairment, whereas in Waters and Badcock, the less severe positive symptoms of the AVH group might have preserved them from exhibiting such impairment. However, the small sample size is an important limitation of the present study, and it may be a possible reason for the difference in findings from other published studies with larger sample sizes.

Our findings are in line with Zhang et al. (2008), who assessed voice familiarity in schizophrenia patients with and without AVH (for more details on this study, see the *Introduction*). This study found that in a task in which participants were presented the recorded voices of their own relatives and friends, and they have to classify them as familiar or nonfamiliar, only the AVH group was impaired, not NAVH patients and healthy controls. Taken together, evidence for a voice familiarity deficit in AVH patients solely as shown by Zhang et al., in addition to a voice identity recognition deficit in AVH patients compared with healthy controls as shown in the present study, may suggest a specific link between voice recognition impairment and hallucinations.

The present study did not find any interaction between sex of the speaker and group, which seems in contradiction with the study of Waters and Badcock (2009), who found that schizophrenia patients were particularly impaired in recognition of female voices. However, it is important to note methodological differences between both studies, particularly in the experimental paradigms used. The recollection of voice identity used in the present study exceeds a simple perceptual judgement (*e.g.*, the sex labelling task used by Waters and Badcock, 2009) and involves hierarchically higher memory processes (Damjanovic and Hanley, 2007). Thus, it is possible that an increase in task demands resulted in impaired recognition for both female and male voices in schizophrenia patients of the present study.

Alternative explanations for the present findings should also be addressed. It is reasonable to assume that the AVH group is generally more symptomatic and thus presents a larger cognitive impairment than the NAVH group does. It should be noted, however, that differences between groups in premorbid verbal IQ were not found. Moreover, it has been previously shown that especially negative symptoms and not positive symptoms are associated with neurocognitive decline (Honer et al., 2005; Lewandowski et al., 2011). In addition, all groups did not differ in the hearing threshold test; thus, a simple hearing deficit cannot explain the differences between groups in the experimental task either. Furthermore, all groups differed in medication dosage, with the highest medication dosage in the AVH group. The additional medication analysis revealed that medication did not have a significant effect on voice recognition (*i.e.*, hit rates and A′), although the *p*-value might indicate a trend, in which case the effect of medication in the findings cannot be completely dismissed. Finally, although findings from the literature suggest that intrinsic perceptual limitations of voice processing in schizophrenia may relate to hallucinations in particular (Rossell and Boundy, 2005; Shea et al., 2007; Zhang et al., 2008), in the present study, AVH patients also rated higher in the delusion subscale. Thus, we cannot entirely rule out the contributions of other positive symptoms toward the impairment in voice identity recognition.

The groups did not differ in percentage of FA, which may imply that patients as well as controls did not differ in their bias toward a certain response (famous/nonfamous). The response criterion measure B''_D did not show significant differences between groups either, although there was a trend suggesting that patients might have had a slightly more conservative response bias when performing the task. However, because of the fact that this difference was not significant, interpretations should be made with caution. In contrast to our study, previous research found differences in response criterion between groups. For example, Allen et al. (2004) found a bias toward externalizing self-generated speech in schizophrenia patients, especially for those with positive symptoms. The lack of group differences in the response bias in the present study might be partly due to the different methodological approach. The present study used famous voices instead of recorded self-generated speech. As mentioned in the *Introduction*, one's own voice when spoken sounds deeper and more resonant than when played from an external apparatus owing to the loss of bone-conducted pathway. The externalization bias found in the previous study may have been due to the loss of the bone-conducted pathway when listening to recordings of self-generated speech, and AVH patients might have been particularly sensitive to such a pathway. Given that there is a qualitative difference between the perception of one's own voice when spoken than when reproduced from an external apparatus, it is possible that the recorded voices in Allen et al. study were not perceived as one's own, resulting in a bias toward externalizing. However, the bias would not be expected when listening to externally presented famous voices, as the bone-conducted pathway confound would be absent.

Regardless of whether group differences in the response criterion exist, which in the study of Allen et al. (2004) occurred probably because of the relatively unfamiliar own voice, this study and the present study found group differences with particularly low performance of voice recognition in patients with schizophrenia with positive symptoms. Difficulties in voice perception may not invalidate the role of source misattribution deficits in the formation of hallucination; instead, voice identity recognition deficits might have an additive effect in generating hallucinations. A general deficit of voice identity recognition that might be linked to hallucinations is proposed.

To summarize, the present study shows that schizophrenia patients perform poorly in the recognition of famous voices, with patients experiencing hallucinations being particularly affected. This suggests the existence of a voice identity recognition deficit in schizophrenia, which may be linked with hallucinations.

ACKNOWLEDGMENTS

We thank Dr Charles Fernyhough for his valuable comments on an early version of the manuscript. We are grateful to Bruce Donley, Kirsten White, Rachel Peck, Dr Craig McGill and Dr Jose Garcia for helping with patient recruitment. This work was supported by the Experimental Psychology Society small grant awarded to Prof David Milner.

DISCLOSURE

The authors declare no conflict of interest.

REFERENCES

Aaronson D, Watts B (1987) Extensions of Grier computational formulas for a′ and B″ to below-chance performance. *Psychol Bull*. 102:439–442.

Alba-Ferrara L, Fernyhough C, Weis S, Mitchell R, Hausmann M (2012) Contributions of emotional prosody comprehension deficits to the formation of auditory verbal hallucinations in schizophrenia. *Clin Psychol Rev* 32:244–250.

Allen PP, Johns LC, Fu CH, Broome MR, Vythelingum GN, McGuire PK (2004). Misattribution of external speech in patients with hallucinations and delusions. *Schizophr Res*. 69:277–287.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th ed. text revision) Washington, DC: American Psychiatric Association.

Andreasen NC (1984a) *Modified Scale for the Assessment of Negative Symptoms (SANS)*. Iowa City, IA: University of Iowa.

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

Andreasen NC, Flaum M, Arndt S (1992) The Comprehensive Assessment of Symptoms and History (CASH). An instrument for assessing diagnosis and psychopathology. *Arch Gen Psychiatry.* 49:615–623.

Andreasen NC, Pressler M, Nopoulos P, Miller D, Ho BC (2010). Antipsychotic dose equivalents and dose-years: A standardized method for comparing exposure to different drugs. *Biol Psychiatry.* 67:255–262.

Bach DR, Buxtorf K, Grandjean D, Strik WK (2008) The influence of emotion clarity on emotional prosody identification in paranoid schizophrenia. *Psychol Med*. 39:1–12.

Békésy G (1949) The structure of the middle ear and the hearing of one's own voice by bone conduction. *J Acoust Soc Am*. 21:217–232.

Bentall RP (1990) The illusion of reality: A review and integration of psychological research on hallucinations. *Psychol Bull*. 107:82–95.

Damjanovic L, Hanley JR (2007) Recalling episodic and semantic information about famous faces and voices. *Mem Cognit*. 35:1205–1210.

Donaldson W (1996) The role of decision processes in remembering and knowing. *Mem Cognit*. 24:523–533.

Gardiner JM, Ramponi C, Richardson-Klavehn A (2002) Recognition memory and decision processes: A meta-analysis of remember, know, and guess responses. *Memory*. 10:83–98.

Garrido L, Eisner F, McGettigan C, Stewart L, Sauter D, Hanley JR, Schweinberger SR, Warren JD, Duchaine B (2009) Developmental phonagnosia: A selective deficit of vocal identity recognition. *Neuropsychologia*. 47:123–131.

Haddock G, McCarron J, Tarrier N, Faragher EB (1999) Scales to measure dimensions of hallucinations and delusions: The Psychotic Symptom Rating Scales (PSYRATS). *Psychol Med*. 29:879–889.

Hanley JR, Damjanovic L (2009) It is more difficult to retrieve a familiar person's name and occupation from their voice than from their blurred face. *Memory.* 17:830–839.

Hirnstein M, Hausmann M, Lewald J (2007) Functional cerebral asymmetry in auditory motion perception. *Laterality.* 12:87–99.

Honer WG, Kopala LC, Rabinowitz J (2005) Extrapyramidal symptoms and signs in first-episode, antipsychotic exposed and non-exposed patients with schizophrenia or related psychotic illness. *J Psychopharmacol*. 19:277–285.

Johns LC, Rossell S, Frith C, Ahmad F, Hemsley D, Kuipers E, McGuire PK, (2001) Verbal self-monitoring and auditory verbal hallucinations in patients with schizophrenia. *Psychol Med*. 31:705–715.

Jones SR, Fernyhough C (2007) Neural correlates of inner speech and auditory verbal hallucinations: A critical review and theoretical integration. *Clin Psychol Rev.* 27:140–154.

Keefe RS, Arnold MC, Bayen UJ, McEvoy JP, Wilson WH (2002) Source-monitoring deficits for self-generated stimuli in schizophrenia: Multinomial modeling of data from three sources. *Schizophr Res*. 57:51–67.

Leitman DI, Foxe JJ, Butler PD, Saperstein A, Revheim N, Javitt DC (2005) Sensory contributions to impaired prosodic processing in schizophrenia. *Biol Psychiatry.* 58:56–61.

Leitman DI, Hoptman MJ, Foxe JJ, Saccente E, Wylie GR, Nierenberg J, Jalbrzikowski M, Lim KO, Javitt DC, (2007) The neural substrates of impaired prosodic detection in schizophrenia and its sensorial antecedents. *Am J Psychiatry.* 164:474–482.

Leitman DI, Wolf DH, Laukka P, Ragland JD, Valdez JN, Turetsky BI, Gur RE, Gur RC (2011) Not pitch perfect: Sensory contributions to affective communication impairment in schizophrenia. *Biol Psychiatry.* 70:611–618.

Lewandowski KE, Cohen BM, Keshavan MS, Ongur D (2011) Relationship of neurocognitive deficits to diagnosis and symptoms across affective and non-affective psychoses. *Schizophr Res*. 133:212–217.

Matsumoto K, Samson GT, O'Daly OD, Tracy DK, Patel AD, Shergill SS (2006) Prosodic discrimination in patients with schizophrenia. *Br J Psychiatry.* 189:180–181.

Maylor EA (1995) Remembering versus knowing television theme tunes in middle-aged and elderly adults. *Br J Psychol*. 86:21–25.

Morrison G, Sharkey V, Allardyce J, Kelly RC, McCreadie RG (2000) Nithsdale schizophrenia surveys 21: A longitudinal study of National Adult Reading Test stability. *Psychol Med*. 30:717–720.

Nelson HE, Willson J (1991) *The Revised National Adult Reading Test Manual*. Windsor, UK: NFER-Nelson.

Perrot P, Aversano G, Chollet G (2007) Voice disguise and automatic detection: Review and perspectives. *Prog Nonlinear Speech Process*. 4391:101–117, 269.

Rossell SL, Boundy CL (2005) Are auditory-verbal hallucinations associated with auditory affective processing deficits? *Schizophr Res*. 78:95–106.

Shea TL, Sergejew AA, Burnham D, Jones C, Rossell SL, Copolov DL, Egan GF (2007) Emotional prosodic processing in auditory hallucinations. *Schizophr Res*. 90:214–220.

Snodgrass JG, Corwin J (1988) Perceptual identification thresholds for 150 fragmented pictures from the Snodgrass and Vanderwart picture set. *Percept Mot Skills*. 67:3–36.

Stephane M, Thuras P, Nasrallah H, Georgopoulos AP (2003) The internal structure of the phenomenology of auditory verbal hallucinations. *Schizophr Res*. 61:185–193.

Waters FA, Badcock JC (2009) Memory for speech and voice identity in schizophrenia. *J Nerv Ment Dis*. 197:887–891.

Waters FA, Badcock JC (2010) First-rank symptoms in schizophrenia: Reexamining mechanisms of self-recognition. *Schizophr Bull*. 36:510–517.

Waters FA, Woodward T, Allen P, Aleman A, Sommer I (2010) Self-recognition deficits in schizophrenia patients with auditory hallucinations: A meta-analysis of the literature. *Schizophr Bull*. 2012;38:741–750.

Weiss AP, Goff DC, Duff M, Roffman JL, Schacter DL (2008) Distinguishing familiarity-based from source-based memory performance in patients with schizophrenia. *Schizophr Res*. 99:208–217.

Zhang ZJ, Hao GF, Shi JB, Mou XD, Yao ZJ, Chen N (2008) Investigation of the neural substrates of voice recognition in Chinese schizophrenic patients with auditory verbal hallucinations: An event-related functional MRI study. *Acta Psychiatr Scand*. 118:272–280.