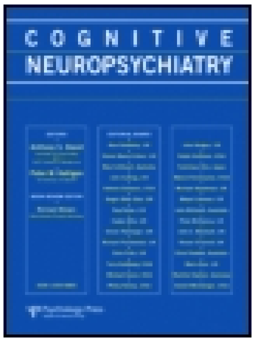


Cognitive Neuropsychiatry

ISSN: 1354-6805 (Print) 1464-0619 (Online) Journal homepage: http://www.tandfonline.com/loi/pcnp20

Is this my voice or yours? The role of emotion and acoustic quality in self-other voice discrimination in schizophrenia

Ana P. Pinheiro, Neguine Rezaii, Andréia Rauber & Margaret Niznikiewicz

COGNITIVE NEUROPSYCHIATRY, 2016
http://dx.doi.org/10.1080/13546805.2016.1208611

Routledge
Taylor & Francis Group

Is this my voice or yours? The role of emotion and acoustic quality in self-other voice discrimination in schizophrenia

Ana P. Pinheiro[a,b,c], Neguine Rezaii[a], Andréia Rauber[d] and Margaret Niznikiewicz[a]

[a]Laboratory of Neuroscience, Department of Psychiatry, Harvard Medical School, & Veterans Affairs Boston Healthcare System, Brockton V.A. Medical Center Psychiatry, Brockton, MA, USA; [b]Neuropsychophysiology Laboratory, CIPsi, School of Psychology, University of Minho, Braga, Portugal; [c]Faculty of Psychology, University of Lisbon, Lisbon, Portugal; [d]Computational Linguistics Department, University of Tübingen, Tübingen, Germany

ABSTRACT

Introduction: Impairments in self-other voice discrimination have been consistently reported in schizophrenia, and associated with the severity of auditory verbal hallucinations (AVHs). This study probed the interactions between voice identity, voice acoustic quality, and semantic valence in a self-other voice discrimination task in schizophrenia patients compared with healthy subjects. The relationship between voice identity discrimination and AVH severity was also explored.

Methods: Seventeen chronic schizophrenia patients and 19 healthy controls were asked to read aloud a list of adjectives characterised by emotional or neutral content. Participants' voice was recorded in the first session. In the behavioural task, 840 spoken words differing in identity (self/non-self), acoustic quality (undistorted/distorted), and semantic valence (negative/positive/neutral) were presented. Participants indicated if the words were spoken in their own voice, another person's voice, or were unsure.

Results: Patients were less accurate than controls in the recognition of self-generated speech with negative content only. Impaired recognition of negative self-generated speech was associated with AVH severity ("voices conversing").

Conclusions: These results suggest that abnormalities in higher order processes (evaluation of the salience of a speech stimulus) modulate impaired self-other voice discrimination in schizophrenia. Abnormal processing of negative self-generated speech may play a role in the experience of AVH.

ARTICLE HISTORY
Received 2 September 2015
Accepted 29 June 2016

KEYWORDS
Schizophrenia; auditory hallucinations; voice; self; emotion

Introduction

Voices plausibly represent the most important sound category in a social environment (e.g., Belin, Fecteau, & Bédard, 2004). Studies in the last decades demonstrated that distinct types of vocal information—identity, emotion, speech—are processed in partially dissociable pathways (e.g., Belin, Bestelmeyer, Latinus, & Watson, 2011; Belin et al., 2004).

Alterations in the neuro-functional mechanisms underlying voice perception have been associated with clinical symptoms such as auditory hallucinations (e.g., Allen et al., 2004;

CONTACT Ana P. Pinheiro ana.pinheiro@psi.uminho.pt

Ford & Mathalon, 2005; Conde, Gonçalves, & Pinheiro, 2016). Representing perceptions without an external stimulus, auditory hallucinations are experienced by up to 70% of schizophrenia patients more often as voices (i.e., auditory verbal hallucinations—AVHs; David, 1999; Mueser, Bellack, & Brady, 1990; Nayani & David, 1996). Similar to the experience of listening to an external voice, hallucinated voices convey information about the identity and emotional state of the speaker, which indicates that AVH engage similar mechanisms to those involved in human voice perception (Badcock & Chhabra, 2013). Therefore, it is likely that specific or general abnormalities in voice information processing pathways (Badcock, 2010; Belin et al., 2004) are associated with the experience of AVH (see Conde et al., 2016 for a review; see also Koeda et al., 2006; Koeda, Takahashi, Matsuura, Asai, & Okubo, 2013; Rapin et al., 2012).

Abnormalities in the processing of voice *identity* in schizophrenia have been consistently reported (e.g., Alba-Ferrara, Weis, Damjanovic, Rowett, & Hausmann, 2012; Badcock & Chhabra, 2013; Kumari et al., 2010; Mou et al., 2013; Stephane, Kuskowski, McClannahan, Surerus, & Nelson, 2010; Waters & Badcock, 2009), and proposed to underlie hallucinatory experiences (Allen et al., 2004; Brookwell, Bentall, & Varese, 2013; Mechelli et al., 2007; Zhang et al., 2008). Schizophrenia patients with hallucinations are more likely than patients without hallucinations and a healthy control group to misidentify self-generated speech (SGS) as externally generated (Allen et al., 2004; Stephane et al., 2010). Misattribution errors may be dependent on the acoustic quality of recorded speech: Allen et al. (2004) found that these errors occurred when patients listened to distorted (i.e., pitch-altered) words, and were positively associated with the severity of hallucinations.

Speech and vocal *emotion* perception abnormalities were also proposed to account for the AVH experience (e.g., Shea et al., 2007; Vercammen, de Haan, & Aleman, 2008; Woodruff et al., 1997). In particular, there is evidence that schizophrenia patients process emotional information in an anomalous way, and that these impairments are associated with AVH severity. For example, when presented with auditory emotional words, increased activation in temporal regions, orbitofrontal cortex, insula, cingulate cortex, and right amygdala was observed in patients with chronic AVH compared with healthy controls (Sanjuan et al., 2007), and in bilateral amygdala and parahippocampal gyrus in the same group compared with both healthy controls and patients without AVH (Escartí et al., 2010). Furthermore, impaired pitch discrimination of pure tones (McLachlan, Phillips, Rossell, & Wilson, 2013), as well as the detection of emotional salience from prosodic cues (Alba-Ferrara et al., 2012; Rossell & Boundy, 2005; Shea et al., 2007), seem to be specifically associated with AVH severity, substantiating the contribution of emotional disturbances in AVH.

Nonetheless, few studies examined the interactions between voice identity and emotion, and their contribution to the experience of AVH. Studies with healthy subjects demonstrate that emotion and self-relevance interactively influence one another during word comprehension (Fields & Kuperberg, 2012; Pinheiro et al., 2016; Watson, Dritschel, Obonsawin, & Jentzsch, 2007), and that self-relevance and emotion exert similar effects (e.g., Fields & Kuperberg, 2012). Studies with schizophrenia patients indicated important interactions between voice identity and semantic valence[1] during voice discrimination tasks, and provided evidence for a negative perceptual bias in patients with AVH. For example, when producing speech and hearing a distorted version of their own voice, schizophrenia patients experiencing AVH, but not healthy controls, consider more often that their own distorted voice is an unfamiliar voice if the words being uttered are *negative* (Johns et al., 2001). External misattribution errors were increased for *negative* affective stimuli (words such as "unfortunate" or "contaminated") in schizophrenia patients with AVH relative to both healthy subjects and patients in remission (Allen et al., 2004[2]; Costafreda, Brébion, Allen, McGuire, & Fu, 2008; Johns et al., 2001). This negative bias may explain why voices often have a negative or frightening tone (e.g., Nayani & David, 1996).

Nonetheless, further study is needed to better understand to what extent speech valence may affect self-other speech discrimination in schizophrenia. Additionally, it remains to be clarified how the acoustic quality of the voice stimuli (undistorted vs. distorted) affects self-other speech discrimination, and whether the negative bias is specific to auditory hallucinations or underlies positive symptoms in general. Previous research has been hampered by the lack of a neutral word condition or by the insufficient number of experimental stimuli that renders conclusions tenuous. Therefore, based on a thoroughly controlled experimental design, with a high number of stimuli per condition, the current study aimed to probe whether voice identity (self vs. other), voice acoustic quality (undistorted vs. distorted), and speech semantic valence (neutral vs. positive vs. negative) interactively modulate the processing of speech information, and how this influences recognition performance in schizophrenia patients. The focus of the analysis was on accuracy and number of "unsure" responses. The relationship between the capacity to discriminate speech stimuli based on voice identity and the severity of auditory hallucinations was also explored.

Considering that schizophrenia patients are impaired in the discrimination between SGS and non-self speech (NSS) stimuli, we expected lower accuracy rates in the recognition of self-generated words relative to words spoken by an unfamiliar individual. Moreover, if semantic valence plays a role in the discrimination between SGS and NSS, we hypothesised that the discrimination difficulties would be enhanced for speech with negative valence relative to neutral and positive words, in line with a negative bias in misattribution errors. Finally, if these impairments are enhanced in patients with AVH, we hypothesised that faulty self-other discrimination in speech of negative valence would be associated with the severity of AVH (e.g., Allen et al., 2004; Costafreda et al., 2008; Johns et al., 2001).

Methods

Participants

Seventeen patients with chronic schizophrenia and 18 healthy controls participated in the behavioural experiment (Table 1). Common inclusion criteria were: American English as native language; age between 18 and 50 years; no history of neurological illness; no history of alcohol or drug dependence in the past five years or abuse within the last year (American Psychiatric Association, 2000); verbal intelligence quotient (IQ) above 85 (Wechsler, 1997); no alcohol use in the 24 hours before testing; no hearing, vision, or upper body impairment. For healthy controls, an additional exclusion criterion was a history of psychiatric disorder in oneself or in first-degree relatives.

Patients were recruited through Internet advertisements, from outpatient clinics, or through referrals from clinicians. They were stabilised outpatients, with no current mood episodes. Patients' diagnoses were based on interviews with the Structured Clinical

Table 1. Demographic, cognitive, and clinical characteristics of patients with schizophrenia and healthy controls.

Variable	Schizophrenia patients ($N = 17$)	Healthy controls ($N = 18$)	t, p[a]
Sociodemographic			
Age	48.29 (8.49)	49.53 (4.48)	.479, .636
Sex	12 males; 5 females	13 males; 5 females	
Education (years)	13.69 (2.33)	15.06 (1.91)	1.909, .065
Parental SES	2.60 (1.45)	2.42 (1.00)	−.372, .713
Cognitive			
General IQ[b]	99.18 (12.34)	108.76 (17.46)	3.585, .001*
Verbal Comprehension Index[b]	98.56 (12.92)	106.44 (17.34)	1.095, .290
Clinical			
Age of onset	29.00 (10.95)	n.a.	n.a.
Duration of illness	21 (12)	n.a.	n.a.
CPZ equivalent	402.41 (338.69)	n.a.	n.a.
PANSS Positive[c]	29.74 (17.88)	n.a.	n.a.
PANSS Negative[c]	29.84 (17.22)	n.a.	n.a.
PANSS General[c]	45.16 (16.30)	n.a.	n.a.
PANSS Total[c]	84.68 (29.19)	n.a.	n.a.
Global SANS[d]	10.84 (5.92)	n.a.	n.a.
SAPS Auditory Hallucinations[e]	3.44 (1.86)	n.a.	n.a.
SAPS Voices commenting[e]	2.63 (2.06)	n.a.	n.a.
SAPS Voices conversing[e]	2.25 (2.08)	n.a.	n.a.
SAPS Global hallucinations[e]	3.63 (1.63)	n.a.	n.a.
SAPS Global delusions[e]	4.13 (1.31)	n.a.	n.a.
Global SAPS[e]	10.50 (3.31)	n.a.	n.a.

Notes: n.a., not applicable; CPZ, chlorpromazine.
[a]Independent samples *t*-test.
[b]WAIS: Wechsler (1997).
[c]PANSS: Kay et al. (1987).
[d]SANS: Andreasen (1983).
[e]SAPS: Andreasen (1984).
*$p < .05$.

Interview for DSM-IV for Axis I (First, Spitzer, Gibbon, & Williams, 2002) and Axis II (First, Gibbon, & Spitzer, 1997). In addition, patients were assessed with the *Positive and Negative Syndrome Scale* (PANNS: Kay, Fiszbein, & Opler, 1987), the *Scale for the Assessment of Positive Symptoms* (SAPS: Andreasen, 1984), and the *Scale for the Assessment of Negative Symptoms* (SANS: Andreasen, 1983). Healthy controls were recruited through Internet advertisements and matched to the patients on age, sex, parental socio-economic status.

In compliance with the Declaration of Helsinki, participants were provided a clear explanation of the study procedure before the experiment, and provided written informed consent for the experimental protocol approved by the local Ethical Committee. Subjects were paid for their participation in the study.

Stimuli

Stimuli included adjectives with neutral (70), positive (70), and negative (70) semantic valence (Table 2 and Appendix 1). Words were controlled for frequency (Brown verbal frequency), concreteness, familiarity, imageability, number of letters, and number of syllables (Coltheart, 2007; Wilson, 1988); the number of phonemes differed between the stimuli types (see Table 2). Valence and arousal ratings were obtained through the norms for 13,915 English lemmas published by Warriner, Kuperman, and Brysbaert (2013) (Appendix 1).

Table 2. Psycholinguistic and affective properties of the words included in the experiment, for each valence type.

	Word valence category			
	Neutral	Positive	Negative	*F*, *p*[a]
Psycholinguistic properties				
Brown verbal frequency[b]	12.61 (24.15)	17.97 (53.51)	18.44 (100.57)	0.162, .850
Concreteness (100–700)[c]	165.34 (202.17)	106.87 (161.07)	111.04 (162.19)	2.400, .093
Familiarity (100–700)[d]	360.71 (32.64)	329.36 (281.16)	257.70 (276.35)	2.546, .081
Imageability[e]	263.76 (218.77)	248.33 (214.50)	200.59 (216.34)	1.619, .201
Number of phonemes	4.43 (1.81)	3.76 (2.89)	3.30 (2.37)	3.921, .021
Number of syllables	1.67 (0.72)	1.81 (0.97)	1.61 (0.82)	1.049, .352
Affective properties				
Valence[f]	5.64 (0.58)	7.34 (0.51)	2.83 (0.57)	710.108, <.001

[a]One-way ANOVA.
[b]This measure includes 14,529 entries that range from 0 to 6,833 ($M = 35$; SD = 252).
[c]Values range from 100 to 700 ($M = 438$; SD = 120).
[d]Values range from 100 to 700 ($M = 488$; SD = 99).
[e]Values range from 100 to 700 ($M = 450$; SD = 108).
[f]Values range from 1 to 9 (Warriner et al., 2013).

Procedure

Before the experimental session, each participant's voice was recorded. Participants were asked to read aloud the list of 210 adjectives (SGS condition) (Appendix 1). The words were shown in the centre of a computer screen, one at a time. Before seeing the word, participants were instructed to listen to that same word pronounced by an unknown speaker using neutral prosody, and afterwards to match the loudness and neutral prosody of each target word at constant voice intensity (65 dB). The inclusion of a "voice-model" aimed at reducing between-subjects variability in speech rate, voice loudness and pitch. Recordings were made in a quiet room with an Edirol R-09 recorder and a CS-15 cardioid-type stereo microphone, with a sampling rate of 44,100 kHz and 16-bit quantisation. For the NSS condition, the same 210 words were recorded by a male (age = 43) or female (age = 44) native speaker of American English unknown to the participants (for a male participant, a male control voice was used; for a female participant, a female control voice was used), using the same procedure described for the SGS. The same "voice-model" was used during SGS and NSS recordings.

After the recording session, each word was segmented using Praat software. First, voice stimuli were normalised according to peak amplitude by means of a Praat script. Audacity software was used for noise reduction. Then, stimuli were cut at the beginning and at the end of each word, using Praat software. Mean pitch, intensity and duration were subsequently calculated for each condition (Table 3). The differences in the acoustic features of SGS and NSS for female and male participants, based on one-sample *t*-tests, are presented in Table 3. We also ran a repeated-measures ANOVA to probe the effects of valence on the acoustic features of the speech stimuli in both groups. The ANOVA included the within-subject factors of valence (neutral, positive, negative) and acoustic feature (duration, minimum F0, mean F0, maximum F0, mean intensity), and group as between-subjects factor. No group differences were observed ($p > .05$). A significant interaction between valence and acoustic feature type ($F(8, 264) = 45.348$, $p < .001$, *partial* $\eta^2 = .579$) indicated valence-related differences in duration and minimum F0. Planned pairwise comparisons revealed that positive words had longer duration than

Table 3. Acoustic properties of the recorded words spoken by healthy controls (SGS condition), schizophrenia patients (SGS condition), and unfamiliar subjects (NSS condition).

Valence	Acoustic measure	SGS		NSS		Comparison between SGS and NSS			
		HC	SZ	NSS-Male	NSS-Female	SGS vs. NSS—Male HC	SGS vs. NSS—Male SZ	SGS vs. NSS—Female HC	SGS vs. NSS—Female SZ
Neutral	DUR (ms)	594.20 (57.03)	641.78 (68.85)	583.50	654.75	.195	2.907*	−2.383	−.413
	$F0_{Min}$ (Hz)	108.45 (23.79)	111.48 (18.27)	75.19	180.94	4.379**	7.710***	−9.208**	−7.350**
	$F0_{M}$ (Hz)	137.82 (29.98)	148.41 (24.48)	110.24	206.58	2.560*	5.047***	−5.202**	−3.141*
	$F0_{Max}$ (Hz)	176.61 (42.96)	191.89 (35.81)	159.90	248.77	.3222	2.187	−2.996*	−1.945
	INT_{M} (dB)	72.74 (3.03)	74.94 (2.38)	72.66	75.86	−.688	2.496*	−1.642	.056
Positive	DUR (ms)	645.26 (66.43)	683.98 (59.92)	628.79	721.70	.443	2.973*	−2.717	−1.434
	$F0_{Min}$ (Hz)	114.66 (37.80)	108.27 (17.44)	77.79	178.97	2.730*	6.336***	−18.067***	−8.223**
	$F0_{M}$ (Hz)	135.78 (30.16)	148.25 (24.34)	101.18	211.97	3.610**	6.263***	−9.467**	−4.222*
	$F0_{Max}$ (Hz)	172.30 (41.66)	194.09 (34.17)	136.07	260.77	2.389*	5.008***	−9.690**	−3.218*
	INT_{M} (dB)	72.21 (2.85)	74.38 (2.46)	71.97	75.34	−.414	2.679*	−1.785	−1.477
Negative	DUR (ms)	637.03 (62.88)	673.49 (64.63)	611.65	703.02	1.036	3.070*	−3.003*	−1.407
	$F0_{Min}$ (Hz)	107.02 (22.69)	108.03 (17.03)	79.24	173.59	3.728**	6.448***	−9.468**	−6.933**
	$F0_{M}$ (Hz)	138.14 (29.81)	147.47 (22.87)	117.89	206.59	1.457	3.997**	−7.671**	−3.585*
	$F0_{Max}$ (Hz)	176.93 (41.90)	190.85 (31.54)	166.37	248.12	−.448	1.732	−7.493**	−1.663
	INT_{M} (dB)	76.92 (19.01)	74.26 (2.32)	72.19	75.92	.936	2.159	−1.743	−2.827*

Notes: HC, healthy controls; SZ, schizophrenia patients; SGS, self-generated speech; NSS, non-self speech; $F0_{Min}$, minimum F0; $F0_{M}$, mean F0; $F0_{Max}$, maximum F0.

*$p < .05$.

**$p < .01$.

***$p < .001$.

neutral ($p < .001$) and negative ($p = .001$) words, and that negative words had longer duration than neutral words ($p < .001$). In addition, minimum F0 was higher in neutral relative to negative words ($p < .001$).

Words in both conditions (SGS, NSS) were additionally pitch-shifted and presented in the *distorted* condition. Following previous studies (e.g., Allen, Freeman, Johns, & McGuire, 2006), the degree of pitch distortion was −4 semitones, since it made the speaker's voice harder to recognise but without the speech becoming incomprehensible.[3] Therefore, the experiment included 840 stimuli (210 undistorted SGS; 210 undistorted NSS; 210 distorted SGS; 210 distorted NSS). To reduce the effects of fatigue, the six combinations of identity and semantic valence were ordered pseudo-randomly, with the constraint of no more than three consecutive trials of the same condition. Words were presented in two lists of 420 words each, and the lists were counterbalanced across participants. Half of the participants received the lists in an AB sequence, and half in BA sequence.

Each participant was seated comfortably at a distance of 100 cm from a computer monitor in a sound-attenuating chamber. Participants indicated if the words were spoken in their own voice, another person's voice, or were unsure, via a button press on a Cedrus response pad (RB-830, San Pedro, CA, USA). The availability of an "unsure" option allowed participants to make a choice between "self" and "other" with some degree of confidence, instead of a forced choice. Order of buttons was counterbalanced.

Before each word onset, a fixation cross was presented centrally on the screen for 1500 ms, and was kept during word presentation to minimise eye movements. After a 1000 ms inter-stimulus interval, a question mark signalled the beginning of the response time (6 s). Stimuli were presented binaurally through headphones at a sound level comfortable for each subject, and were not repeated during the experiment. Stimulus presentation and timing of events and recording of subjects' responses were controlled by Superlab Pro software package (2008; http://www.superlab.com/). Before each experimental block, participants were given a brief training with feedback.

Statistical analyses

Task performance was analysed with respect to accuracy and number of unsure responses. Accuracy was measured as the number of correct responses corrected for false alarms (hits-false alarms; false alarm rates represent the number of times a participant responded "self" when the stimulus was a non-self voice): SGS stimuli accurately identified as "self" and NSS stimuli identified as "other" were considered hits. The number of "unsure" responses was computed for both SGS and NSS stimuli, as a function of semantic valence and acoustic quality, and was analysed separately.

The effects of identity, semantic valence, and acoustic quality on the number of correct responses were tested using repeated-measures ANOVA with identity (SGS, NSS), acoustic quality (undistorted, distorted), and valence (neutral, positive, negative) as within-subject factors, and group as between-subjects factor. The same statistical model was applied to the number of "unsure" responses. In order to rule out the effects of differences in the acoustic properties of vocal stimuli, the mean differences in duration, F0, and intensity between SGS and NSS were added as covariates. An additional ANOVA explored the effects of education and full scale IQ on the behavioural performance, adding these two variables as covariates.

Analyses were corrected for non-sphericity using the Greenhouse–Geisser method (the original df is reported). Significant interactions were followed with pairwise comparisons between conditions, using the Bonferroni adjustment for multiple comparisons. Effect sizes for significant effects are reported using the partial η-square method (*partial* η^2).

Results

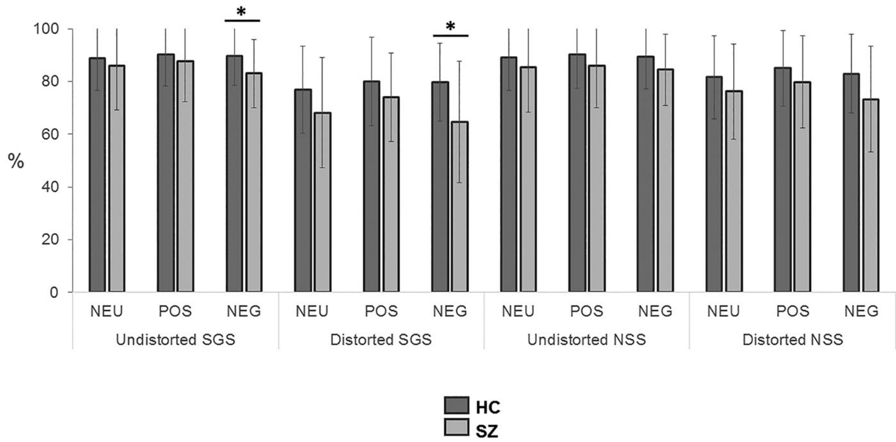
Accuracy and unsure responses

Accuracy

None of the covariates was significantly related to recognition accuracy, including education and IQ ($p > .05$).[4] Effects that were similar for both groups included identity, semantic valence, and the interaction between voice identity and acoustic quality, which were significant after controlling for the effect of voice acoustic properties. NSS stimuli were more often accurately recognised as "other" than SGS stimuli were recognised as "self", as indicated by a significant main effect of voice identity ($F(1, 30) = 6.478$, $p = .016$, *partial* $\eta^2 = .178$). A significant identity by acoustic quality interaction ($F(1, 30) = 15.333$, $p < .001$, *partial* $\eta^2 = .338$) revealed differences between SGS and NSS in the distorted speech condition, but not in the undistorted speech condition. Planned pairwise comparisons indicated that participants of both groups were more accurate in recognising distorted NSS stimuli as "other" ($p = .001$) than in recognising distorted SGS stimuli as "self" ($p > .05$). Furthermore, participants were more accurate in recognising speech identity ("self" vs. "other") when the words had positive valence than when the words were both neutral ($p < .001$) or negative ($p = .003$) (effect of semantic valence—$F(2, 60) = 10.003$, $p < .001$, *partial* $\eta^2 = .250$). However, no differences in the recognition of neutral and negative speech stimuli were observed ($p > .05$).

Group differences were revealed by a significant group by voice identity by semantic valence interaction after controlling for the effects of voice acoustic properties ($F(2, 60) = 4.615$, $p = .014$, *partial* $\eta^2 = .133$). We followed this interaction by computing separate ANOVAs for SGS and NSS, with valence and distortion as within-subject factors. In the case of SGS, a significant group by emotion interaction was observed ($F(2, 60) = 7.987$, $p = .001$, *partial* $\eta^2 = .210$). Specifically, groups differed in the recognition of SGS with negative content: patients were less accurate than controls ($t(33) = 2.237$, $p = 032$), that is, they misrecognised self-generated words such as "rude" or "ugly" more often as "other" (Figure 1).

In order to rule out whether there were acoustic differences between negative SGS speech correctly recognised as "self" and negative SGS misidentified as "other" by schizophrenia patients, we ran an additional ANOVA on undistorted and distorted self-generated negative speech, adding the difference between self and non-self voices in pitch, intensity and duration for negative speech only as covariates. Even though the group effect remained significant ($F(1, 30) = 4.832$, $p = .036$, *partial* $\eta^2 = .139$), the effects of the covariates were not significant (mean duration—$F(1, 30) = .051$, $p = .823$; mean F0—$F(1, 30) = .367$, $p = .549$; mean intensity—$F(1, 30) = .052$, $p = .821$).[5]



% of False Alarms												
	Undistorted SGS			Distorted SGS			Undistorted NSS			Distorted NSS		
	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative
HC	4.29 (6.88)	3.73 (6.56)	4.52 (7.56)	10.71 (12.55)	8.49 (11.47)	10.24 (10.02)	3.73 (6.52)	3.49 (6.43)	3.17 (4.32)	3.97 (7.50)	3.57 (7.48)	3.97 (7.92)
SZ	3.19 (3.45)	3.28 (3.90)	4.20 (4.39)	13.61 (12.94)	9.41 (11.49)	15.55 (15.62)	3.87 (7.08)	3.61 (5.10)	4.79 (6.16)	3.03 (4.31)	2.77 (4.70)	3.78 (5.82)

Figure 1. Percentage of correct responses, corrected for number of false alarms, according to condition and group. Percentage of false alarms is also shown.
Notes: Error bars represent standard deviations. HC, healthy controls; SZ, schizophrenia; NEU, neutral; POS, positive; NEG, negative; SGS, self-generated speech; NSS, non-self speech; False alarms represent the number of times participants responded "self" when the stimulus was a non-self voice (% values are shown for each experimental condition); *$p < .05$.

Unsure responses

None of the covariates was significantly related to the number of unsure responses, including education and IQ ($p > .05$). A significant interaction between identity, acoustic quality, and semantic valence factors ($F(2, 60) = 16.869$, $p < .001$, *partial* $\eta^2 = .360$) was observed after controlling for the effects of voice acoustic properties (Figure 2). We followed up this interaction by probing the effects of acoustic quality and semantic valence in SGS and NSS separately. A significant effect of distortion ($F(1, 30) = 4.276$, $p = .045$, *partial* $\eta^2 = .127$) was observed when analysing SGS separately: "unsure" responses were increased in the condition of perceptual ambiguity, that is, when speech was distorted, irrespective of valence type.

Overall, patients tended to indicate they were unsure about the identity of the words more often than controls, but this effect was not statistically significant after controlling for the effects of voice acoustic properties (main effect of group–$F(1, 30) = 3.740$, $p = .063$).[6]

Correlations between hallucination scales and accuracy rates

One-tailed Pearson correlation analyses were conducted in an exploratory analysis of the relationship between lower accuracy in the schizophrenia group and hallucination severity, uncorrected for multiple comparisons. The significance threshold was set at .05.

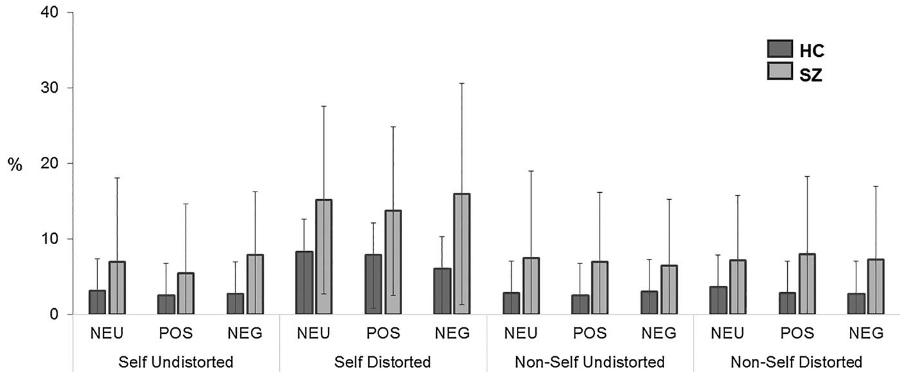


Figure 2. Percentage of unsure responses according to condition and group.
Notes: Error bars represent standard deviations. HC, healthy controls; SZ, schizophrenia; NEU, neutral; POS, positive; NEG, negative.

Accuracy in the recognition of undistorted SGS with negative content was significantly associated with the severity of SAPS "voices conversing" score ($r = -.507$, $p = .023$) only: the higher the SAPS score the lower the accuracy. This item measures the severity of reports of hearing two or more voices conversing on a scale from 0 to 9. No significant association was found between accuracy and other positive symptoms ($p > .05$; see Table 4).

Discussion

This study probed the role of semantic valence and acoustic distortion in self-other voice discrimination in schizophrenia. An additional goal was to explore the association between these processes and the severity of auditory hallucinations. The results highlighted both similarities and differences in voice discrimination in healthy subjects and schizophrenia patients. In both groups, we observed the same tendency for increased misattribution errors for distorted SGS (Allen et al., 2004; Johns et al., 2001). Contrary to other studies, we did not observe that schizophrenia impairments were enhanced under conditions of perceptual ambiguity, that is, when voice stimuli were distorted (e.g., Allen et al., 2004). The tendency for misidentifying distorted SGS as "non-self" in both groups is related to the fact that introducing a pitch change results in a more difficult recognition of SGS as "self" (e.g., Allen et al., 2006). Since an "unsure" option was included, this finding does not seem to reflect a forced choice between "self" and "other", thereby revealing some degree of confidence in the identification response (e.g., Allen et al.,

Table 4. Pearson correlations between accuracy in the recognition of negative SGS and positive symptoms (SAPS) in schizophrenia.

	Auditory hallucinations	Voices commenting	Voices conversing	Global hallucinations	Global delusions	Global SAPS
Undistorted negative SGS	−.230	−.273	−.487* ($p = .028$)	−.138	−.073	−.233
Distorted negative SGS	.150	−.068	−.313	.150	.051	.166

*$p < .05$.

2004). Furthermore, both groups were more accurate in identifying NSS than SGS. These findings provide further support against the idea that misattribution errors are due to a general cognitive impairment in schizophrenia (see also Johns et al., 2001). Furthermore, the lack of group differences in the recognition of NSS suggests that the perceptual voice discrimination failure pointed out by previous studies (e.g., Alba-Ferrara et al., 2012; Zhang et al., 2008) is not an exclusive factor accounting for disrupted self-other voice discrimination in schizophrenia. If that were the case, we would have observed lower recognition for speech stimuli irrespective of emotional category. Instead, top-down factors should be considered, as discussed later.

Impaired attribution of agency is specific of negative speech in schizophrenia

Group differences emerged in the number of correct responses corrected for false alarm rates. In particular, a difference emerged in the identification of SGS with negative content (e.g., "stupid", "abnormal"). Consistent with the literature, external misattributions (recognising SGS as non-self) in schizophrenia were more frequent than self-misattributions (recognising NSS as self-generated). However, the externalising bias reported in patients was only significant in the case of speech with negative content (e.g., "abnormal" or "crazy").

The specific difference in the recognition of negative SGS lends support for a negativity bias in the misattribution errors found during voice discrimination in schizophrenia. Additionally, this finding suggests that schizophrenia patients' impairments in perceiving agency over their own voice (i.e., externalising errors) reported in previous studies (e.g., Allen et al., 2004) are not generalisable to all types of SGS. In line with some of the previous studies, this emotion-specific finding indicates that schizophrenia patients may be particularly prone to attribute negative speech to an external source (e.g., Johns et al., 2001). Previous neuroimaging studies demonstrated that alterations in the structural and functional connectivity between frontal, temporal and cingulate regions might contribute to impairments in self-other voice discrimination (e.g., Allen et al., 2007; Mou et al., 2013). The biased judgment of the source of speech stimuli (self vs. non-self) has been associated with aberrant top-down modulation of temporal regions (e.g., superior temporal gyrus) by the anterior cingulate (Allen et al., 2007). Moreover, the contribution of disturbed fronto-temporal connectivity to reduced accuracy in voice identity recognition was highlighted (Mou et al., 2013).

Impaired attribution of agency to negative self-speech is associated with hallucination severity

The specific impairment in the recognition of SGS with negative content was associated with the severity of auditory hallucinations, in particular the feature of "voices conversing". This finding supports the involvement of externalising biases in hallucinatory experiences (e.g., Brookwell et al., 2013; Varese, Barkus, & Bentall, 2011). The observation that patients with AVH have difficulties in distinguishing internally and externally generated perceptual signals is not a new one (Tian & Poeppel, 2012). Specifically, the association found in our study suggests that attribution processes (self vs. non-self) are biased by emotional factors and may play a role in the aetiology of hallucinations. For example, the content of AVH drives the emotional evaluation of the hallucinatory voice, resulting in high levels of negative affect (e.g., de Leede-Smith & Barkus, 2013). As proposed by Smailes, Moseley, and Wilkinson (2015), negative affect may play a role in AVH development by reducing a sense of agency for self-generated actions (see also Hoskin, Hunter, & Woodruff, 2014; Smailes, Meins, & Fernyhough, 2014).

Alternatively, auditory hallucinations—or an altered sense of "self"—may result in impairments in the processing of speech with emotional valence (Brunet-Gouet & Decety, 2006). A previous study demonstrated that a lack of synchronisation between Broca's area and its homologue in the right hemisphere may lead to an erroneous interpretation of emotional speech from the right hemisphere as coming from an external source (Ćurčić-Blake et al., 2013). Also, schizophrenia patients with AVH may show increased neural activity in response to emotional words specific to the content of the hallucinations (e.g., Sanjuan et al., 2007). This would be manifested in under-attribution of agency ("it is not my voice") when stimuli are congruent with the typical negative content of the hallucinatory voices and match their expectations (see also Nazimek, Hunter, & Woodruff, 2012). Even though the question of what neurocognitive processes underlie the negative emotional tone of AVH experienced by schizophrenia patients has been intriguing researchers for decades, there is not a satisfactory explanation yet. This should be the focus of future studies.

The specific association between impaired negative SGS recognition and the "voices conversing" SAPS item may represent the social nature of AVH ("a community of one", Bell, 2013). Indeed, phenomenological descriptions of the AVH experience suggest that this experience is, for most of the hearers, a social one—often voices are perceived as conversational voices with a social identity, and hearers may interact with them in a way that resembles real social interactions (e.g., Bell, 2013; Woods, Jones, Alderson-Day, Callard, & Fernyhough, 2015). Voices are less commonly described as single words or brief phrases (e.g., Woods et al., 2015). It is also worth noting that a robust body of evidence suggests that different subtypes of AVH should be considered as a function of the underlying phenomenology (e.g., Hugdahl, 2015). Therefore, the association with the "voices conversing" SAPS item and not with more general measures of hallucinatory experience may reflect the heterogeneity of the AVH experience, that is, the conversational aspect of AVH may be more predominant in our group of patients than in others. Future studies should address this possibility.

The emotional valence of the voice seems to represent a critical difference in the experience of AVH in healthy and psychotic voice hearers (e.g., Sommer et al., 2010). We should note that the content of hallucinations is not negative in all voice hearers. Indeed, AVH may be perceived as positive and benevolent (more often) in non-psychotic individuals with high hallucination predisposition (e.g., de Leede-Smith & Barkus, 2013), and there is increasing support for subtypes of AVH (e.g., McCarthy-Jones et al., 2014). As such, our findings should be interpreted in the context of AVH experience in psychotic populations, and are not generalisable to all types of AVH experience.

Bottom-up and top-down factors in the misattribution errors for negative SGS

As the specific group differences reached statistical significance after controlling for the effects of voice acoustic properties (see Stephane et al., 2010 for a similar result), our data speak against the hypothesis that the impaired recognition of SGS with negative content is primarily associated with bottom-up impairments in the use of acoustic cues for voice differentiation.[7] Rather, it provides support for the hypothesis that higher order processes, such as emotional salience attribution or evaluation, play a significant role in altered self-other voice discrimination. This bias is congruent with phenomenological reports of AVH suggesting that voices are often perceived as malevolent, uttered with negative and derogatory content, including words of vilification and abuse (Allen et al., 2004; Larøi et al., 2012; Nayani & David, 1996). In addition, it leaves open the possibility that top-down factors play a bigger role in the generation of AVH (e.g., Nazimek et al., 2012). Indeed, more recent research has substantiated the claim that the form and content of AVH arise through an interaction between hypersalient auditory signals and top-down mechanisms, such as prior knowledge, mental imagery or perceptual expectations (Waters et al., 2012).

Conclusions

This study demonstrated the modulatory effects of semantic valence on altered self-other speech discrimination in schizophrenia. Schizophrenia patients were less accurate in identifying SGS with negative content as "self", and this specific impairment was associated with the severity of auditory hallucinations. The current findings do not support a general perceptual voice discrimination failure in schizophrenia. Instead, they suggest that a failure in self-other voice discrimination may be related to higher order processes such as the evaluation of the emotional relevance of speech and, in particular, of words such as "ugly", "abnormal", or "stupid".

Notes

1. The term "semantic valence" is used throughout the manuscript to indicate neutral vs. emotional semantic content, as opposed to valence associated with prosodic content.
2. In the study of Allen et al. (2004), the tendency for increased misattribution errors in patients with hallucinations when the words had negative content (e.g., "corrupt", "contaminated", "unfortunate") was not statistically significant. As the number of trials per condition was relatively low (24 undistorted SGS, 24 distorted SGS, 24 undistorted non-self speech, 24 distorted non-self speech, with approximately only 8 negative, 8 positive, and 8 neutral words in each condition), this may have resulted in a lack of power reflected in the lack of statistically significant differences related to the negative speech condition.
3. Lowering the pitch, instead of increasing it, was justified by previous evidence indicating that it yields more prominent neural responses to pitch feedback perturbations (Chen et al., 2012; Liu, Meshman, Behroozmand, & Larson, 2011).
4. In a complementary analysis, we tested the effects of education and general IQ on the behavioural data, by computing a separate ANOVA with these variables as covariates. The effects of education ($p = .921$) and general IQ ($p = .723$) were not significant.
5. In order to rule out whether there were acoustic differences between negative SGS speech correctly recognized as "self" and negative SGS misidentified as "other" by schizophrenia patients, we also performed an acoustic analysis of these speech stimuli, based on the individual performance of each patient. The paired samples t-tests did not reveal differences in the acoustic properties of negative SGS associated with correct responses and with errors (duration – $p = .620$; mean F0 – $p = .710$; mean intensity – $p = .337$).
6. In order to rule out the effects of the number of "unsure" responses, we ran a repeated-measures ANOVA on the proportion of correct responses, adjusted for the number of

"self" and "other" responses in each semantic valence and acoustic quality condition (i.e., the number of "unsure" responses was subtracted from the total number of available responses). A significant group by semantic valence interaction was observed after controlling for the effects of voice acoustic properties ($F(2, 60) = 4.305$, $p = .018$, *partial* $\eta^2 = .125$). We followed up this interaction by running separate ANOVAs for each semantic valence type, keeping identity and acoustic quality as within-subject factors. A significant group by identity by acoustic quality interaction was observed when analysing negative speech ($F(1, 33) = 4.736$, $p = .037$, *partial* $\eta^2 = .126$), confirming a specific impairment in the recognition of the identity of negative SGS in schizophrenia.

7. We note, though, that we did not control for formant dispersion, contrary to previous studies (e.g., Chhabra, Badcock, Maybery, & Leung, 2012). We focus our acoustic analysis on F0, as this measure has been shown to play the most critical role in self-voice recognition and in the discrimination between familiar and unfamiliar voices (e.g., Baumann & Belin, 2010; Latinus & Belin, 2012; Latinus, McAleer, Bestelmeyer, & Belin, 2013; Xu, Homae, Hashimoto, & Hagiwara, 2013).

Acknowledgments

We are grateful to all the participants of this study for their contribution to science.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Grants IF/00334/2012 and PTDC/PSI-PCL/116626/2010 awarded to A.P.P. The two grants were funded by *Fundação para a Ciência e a Tecnologia* (FCT, Portugal) and, in addition, by FEDER (*Fundo Europeu de Desenvolvimento Regional*) through the European programs QREN and COMPETE.

References

Alba-Ferrara, L., Weis, P. S., Damjanovic, P. L., Rowett, M., & Hausmann, M. (2012). Voice identity recognition failure in patients with schizophrenia. *Journal of Nervous and Mental Disease*, *200*(9), 784–790. http://doi.org/10.1097/NMD.0b013e318266f

Allen, P., Amaro, E., Fu, C. H. Y., Williams, S. C. R., Brammer, M. J., Johns, L. C., & McGuire, P. K. (2007). Neural correlates of the misattribution of speech in schizophrenia. *The British Journal of Psychiatry*, *190*, 162–169. http://doi.org/10.1192/bjp.bp.106.025700

Allen, P., Freeman, D., Johns, L., & McGuire, P. (2006). Misattribution of self-generated speech in relation to hallucinatory proneness and delusional ideation in healthy volunteers. *Schizophrenia Research*, *84*(2–3), 281–288. http://doi.org/10.1016/j.schres.2006.01.021

Allen, P., Johns, L. C., Fu, C. H. Y., Broome, M. R., Vythelingum, G. N., & McGuire, P. K. (2004). Misattribution of external speech in patients with hallucinations and delusions. *Schizophrenia Research*, *69*(2–3), 277–287. http://doi.org/10.1016/j.schres.2003.09.008

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders 4th edition TR*. Washington, DC: American Psychiatric Association Press. http://doi.org/10.1016/B978-1-4377-2242-0.00016-X

Andreasen, N. C. (1983). *Scale for the assessment of negative symptoms*. Iowa City, IA: University of Iowa.

Andreasen, N. C. (1984). *Scale for the assessment of positive symptoms*. Iowa City, IA: University of Iowa.

Badcock, J. C. (2010). The cognitive neuropsychology of auditory hallucinations: A parallel auditory pathways framework. *Schizophrenia Bulletin*, *36*(3), 576–584. http://doi.org/10.1093/schbul/sbn128

Badcock, J. C., & Chhabra, S. (2013). Voices to reckon with: Perceptions of voice identity in clinical and non-clinical voice hearers. *Frontiers in Human Neuroscience*, *7*, 114. http://doi.org/10.3389/fnhum.2013.00114

Baumann, O., & Belin, P. (2010). Perceptual scaling of voice identity: Common dimensions for different vowels and speakers. *Psychological Research*, *74*(1), 110–120. http://doi.org/10.1007/s00426-008-0185-z

Belin, P., Bestelmeyer, P. E. G., Latinus, M., & Watson, R. (2011). Understanding voice perception. *British Journal of Psychology*, *102*(4), 711–725. http://doi.org/10.1111/j.2044-8295.2011.02041.x

Belin, P., Fecteau, S., & Bédard, C. (2004). Thinking the voice: Neural correlates of voice perception. *Trends in Cognitive Sciences*, *8*(3), 129–135. http://doi.org/10.1016/j.tics.2004.01.008

Bell, V. (2013). A community of one: Social cognition and auditory verbal hallucinations. *PLoS Biology*, *11*(12), 1–4. http://doi.org/10.1371/journal.pbio.1001723

Brookwell, M. L., Bentall, R. P., & Varese, F. (2013). Externalizing biases and hallucinations in source-monitoring, self-monitoring and signal detection studies: A meta-analytic review. *Psychological Medicine*, *43*(12), 2465–2475. http://doi.org/10.1017/S0033291712002760

Brunet-Gouet, E., & Decety, J. (2006). Social brain dysfunctions in schizophrenia: A review of neuroimaging studies. *Psychiatry Research – Neuroimaging*, *148*(2–3), 75–92. http://doi.org/10.1016/j.pscychresns.2006.05.001

Chhabra, S., Badcock, J. C., Maybery, M. T., & Leung, D. (2012). Voice identity discrimination in schizophrenia. *Neuropsychologia, 50*(12), 2730–2735. http://doi.org/10.1016/j.neuropsychologia.2012.08.006

Chen, Z., Liu, P., Wang, E. Q., Larson, C. R., Huang, D., & Liu, H. (2012). ERP correlates of language-specific processing of auditory pitch feedback during self-vocalization. *Brain and Language*, *121*(1), 25–34. http://doi.org/10.1016/j.bandl.2012.02.004

Coltheart, M. (2007). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology Section A*, *33*(4), 497–505. http://doi.org/10.1080/14640748108400805

Conde, T., Gonçalves, O. F., & Pinheiro, A. P. (2016). A cognitive neuroscience view of voice processing abnormalities in schizophrenia: A window into auditory verbal hallucinations? *Harvard Review of Psychiatry*, *24*(2), 148–163. http://doi.org/10.1097/HRP.0000000000000082

Costafreda, S. G., Brébion, G., Allen, P., McGuire, P. K., & Fu, C. H. Y. (2008). Affective modulation of external misattribution bias in source monitoring in schizophrenia. *Psychological Medicine*, *38*(6), 821–824. http://doi.org/10.1017/S0033291708003243

Ćurčić-Blake, B., Liemburg, E., Vercammen, A., Swart, M., Knegtering, H., Bruggeman, R., & Aleman, A. (2013). When Broca goes uninformed: Reduced information flow to Broca's area in schizophrenia patients with auditory hallucinations. *Schizophrenia Bulletin*, *39*(5), 1087–1095. http://doi.org/10.1093/schbul/sbs107

David, A. S. (1999). Auditory hallucinations: Phenomenology, neuropsychology and neuroimaging update. *Acta Psychiatrica Scandinavica. Supplementum*, *395*, 95–104. http://doi.org/10.1111/j.1600-0447.1999.tb05988.x

Escartí, M. J., de la Iglesia-Vayá, M., Martí-Bonmatí, L., Robles, M., Carbonell, J., Lull, J. J., … Sanjuán, J. (2010). Increased amygdala and parahippocampal gyrus activation in schizophrenic patients with auditory hallucinations: An fMRI study using independent component analysis. *Schizophrenia Research*, *117*(1), 31–41. http://doi.org/10.1016/j.schres.2009.12.028

Fields, E. C., & Kuperberg, G. R. (2012). It's all about you: An ERP study of emotion and self-relevance in discourse. *NeuroImage*, *62*(1), 562–574. http://doi.org/10.1016/j.neuroimage.2012.05.003

First, M. B., Gibbon, M., & Spitzer, R. (1997). *Structured clinical interview for DSM-IV Axis II personality disorders (SCID-II, version 2.0)*. New York, NY: Biometrics Research Department, New York State Psychiatric Institute.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (2002). *Structured clinical interview for DSM-IV Axis I Diagnosis-Patient Edition (SCID-I/P, version 2.0)*. New York, NY: Biometric Research Department, New York State Psychiatric Institute. http://doi.org/M
Ford, J. M., & Mathalon, D. H. (2005). Corollary discharge dysfunction in schizophrenia: Can it explain auditory hallucinations? *International Journal of Psychophysiology*, *58*(2–3), 179–189. http://doi.org/10.1016/j.ijpsycho.2005.01.014
Hoskin, R., Hunter, M. D., & Woodruff, P. W. R. (2014). The effect of psychological stress and expectation on auditory perception: A signal detection analysis. *British Journal of Psychology*, *105*(4), 524–546. http://doi.org/10.1111/bjop.12048
Hugdahl, K. (2015). Auditory hallucinations: A review of the ERC "VOICE" project. *World Journal of Psychiatry*, *5*(2), 193–209. http://doi.org/10.5498/wjp.v5.i2.193
Johns, L. C., Rossell, S., Frith, C., Ahmad, F., Hemsley, D., Kuipers, E., & McGuire, P. K. (2001). Verbal self-monitoring and auditory verbal hallucinations in patients with schizophrenia. *Psychological Medicine*, *31*, 705–715. http://doi.org/10.1017/S0033291701003774
Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*, *13*(2), 261–276. http://doi.org/10.1093/schbul/13.2.261
Koeda, M., Takahashi, H., Matsuura, M., Asai, K., & Okubo, Y. (2013). Cerebral responses to vocal attractiveness and auditory hallucinations in schizophrenia: A functional MRI study. *Frontiers in Human Neuroscience*, *7*, 221. http://doi.org/10.3389/fnhum.2013.00221
Koeda, M., Takahashi, H., Yahata, N., Matsuura, M., Asai, K., Okubo, Y., & Tanaka, H. (2006). Language processing and human voice perception in schizophrenia: A functional magnetic resonance imaging study. *Biological Psychiatry*, *59*, 948–957. http://doi.org/10.1016/j.biopsych.2006.01.013
Kumari, V., Fannon, D., Ffytche, D. H., Raveendran, V., Antonova, E., Premkumar, P., … Kuipers, E. (2010). Functional MRI of verbal self-monitoring in schizophrenia: Performance and illness-specific effects. *Schizophrenia Bulletin*, *36*(4), 740–755. http://doi.org/10.1093/schbul/sbn148
Larøi, F., Sommer, I. E., Blom, J. D., Fernyhough, C., Ffytche, D. H., Hugdahl, K., … Waters, F. (2012). The characteristic features of auditory verbal hallucinations in clinical and nonclinical groups: State-of-the-art overview and future directions. *Schizophrenia Bulletin*, *38*(4), 724–733. http://doi.org/10.1093/schbul/sbs061
Latinus, M., & Belin, P. (2012). Perceptual auditory aftereffects on voice identity using brief vowel stimuli. *PloS One*, *7*(7), e41384. http://doi.org/10.1371/journal.pone.0041384
Latinus, M., McAleer, P., Bestelmeyer, P. E. G., & Belin, P. (2013). Norm-based coding of voice identity in human auditory cortex. *Current Biology*, *23*(12), 1075–1080. http://doi.org/10.1016/j.cub.2013.04.055
de Leede-Smith, S., & Barkus, E. (2013). A comprehensive review of auditory verbal hallucinations: Lifetime prevalence, correlates and mechanisms in healthy and clinical individuals. *Frontiers in Human Neuroscience*, *7*, 367. http://doi.org/10.3389/fnhum.2013.00367
Liu, H., Meshman, M., Behroozmand, R., & Larson, C. R. (2011). Differential effects of perturbation direction and magnitude on the neural processing of voice pitch feedback. *Clinical Neurophysiology*, *122*(5), 951–957. http://doi.org/10.1016/j.clinph.2010.08.010
McCarthy-Jones, S., Trauer, T., Mackinnon, A., Sims, E., Thomas, N., & Copolov, D. L. (2014). A new phenomenological survey of auditory hallucinations: Evidence for subtypes and implications for theory and practice. *Schizophrenia Bulletin*, *40*(1), 231–235. http://doi.org/10.1093/schbul/sbs156
McLachlan, N. M., Phillips, D. S., Rossell, S. L., & Wilson, S. J. (2013). Auditory processing and hallucinations in schizophrenia. *Schizophrenia Research*, *150*(2–3), 380–385. http://doi.org/10.1016/j.schres.2013.08.039
Mechelli, A., Allen, P., Amaro, E., Fu, C. H. Y., Williams, S. C. R., Brammer, M. J., … McGuire, P. K. (2007). Misattribution of speech and impaired connectivity in patients with auditory verbal hallucinations. *Human Brain Mapping*, *28*(11), 1213–1222. http://doi.org/10.1002/hbm.20341
Mou, X., Bai, F., Xie, C., Shi, J., Yao, Z., Hao, G., … Zhang, Z. (2013). Voice recognition and altered connectivity in schizophrenic patients with auditory hallucinations. *Progress in Neuro-*

Psychopharmacology & Biological Psychiatry, *44*, 265–270. http://doi.org/10.1016/j.pnpbp.2013.03.006
Mueser, K. T., Bellack, A. S., & Brady, E. U. (1990). Hallucinations in schizophrenia. *Acta Psychiatrica Scandinavica*, *82*(1), 26–29.
Nayani, T. H., & David, A. S. (1996). The auditory hallucination: A phenomenological survey. *Psychological Medicine*, *26*(1), 177–189. http://doi.org/10.1017/S003329170003381X
Nazimek, J. M., Hunter, M. D., & Woodruff, P. W. R. (2012). Auditory hallucinations: Expectation–perception model. *Medical Hypotheses*, *78*(6), 802–810. http://doi.org/10.1016/j.mehy.2012.03.014
Pinheiro, A. P., Rezaii, N., Nestor, P., Rauber, A., Spencer, K. M., & Niznikiewicz, M. (2016). Did you or I say pretty, rude or brief? An ERP study of the effects of speaker's identity on emotional word processing. *Brain and Language*, *153–154*, 38–49. http://doi.org/10.1016/j.bandl.2015.12.003
Rapin, L. A., Dohen, M., Lœvenbruck, H., Whitman, J. C., Metzak, P. D., & Woodward, T. S. (2012). Hyperintensity of functional networks involving voice-selective cortical regions during silent thought in schizophrenia. *Psychiatry Research – Neuroimaging*, *202*, 110–117. http://doi.org/10.1016/j.pscychresns.2011.12.014
Rossell, S. L., & Boundy, C. L. (2005). Are auditory-verbal hallucinations associated with auditory affective processing deficits? *Schizophrenia Research*, *78*, 95–106. http://doi.org/10.1016/j.schres.2005.06.002
Sanjuan, J., Lull, J. J., Aguilar, E. J., Martí-Bonmatí, L., Moratal, D., Gonzalez, J. C., … Keshavan, M. S. (2007). Emotional words induce enhanced brain activity in schizophrenic patients with auditory hallucinations. *Psychiatry Research*, *154*(1), 21–29. http://doi.org/10.1016/j.pscychresns.2006.04.011
Shea, T. L., Sergejew, A. A., Burnham, D., Jones, C., Rossell, S. L., Copolov, D. L., & Egan, G. F. (2007). Emotional prosodic processing in auditory hallucinations. *Schizophrenia Research*, *90*(1–3), 214–220. http://doi.org/10.1016/j.schres.2006.09.021
Smailes, D., Meins, E., & Fernyhough, C. (2014). The impact of negative affect on reality discrimination. *Journal of Behavior Therapy and Experimental Psychiatry*, *45*(3), 389–395. http://doi.org/10.1016/j.jbtep.2014.04.001
Smailes, D., Moseley, P., & Wilkinson, S. (2015). A commentary on: Affective coding: The emotional dimension of agency. *Frontiers in Human Neuroscience*, *9*, 142. http://doi.org/10.3389/fnhum.2015.00142
Sommer, I. E., Daalman, K., Rietkerk, T., Diederen, K. M., Bakker, S., Wijkstra, J., & Boks, M. P. M. (2010). Healthy individuals with auditory verbal hallucinations; Who are they? psychiatric assessments of a selected sample of 103 subjects. *Schizophrenia Bulletin*, *36*(3), 633–641. http://doi.org/10.1093/schbul/sbn130
Stephane, M., Kuskowski, M., McClannahan, K., Surerus, C., & Nelson, K. (2010). Evaluation of speech misattribution bias in schizophrenia. *Psychological Medicine*, *40*(5), 741–748. http://doi.org/10.1017/S003329170999081X
Tian, X., & Poeppel, D. (2012). Mental imagery of speech: Linking motor and perceptual systems through internal simulation and estimation. *Frontiers in Human Neuroscience*, *6*, 314. http://doi.org/10.3389/fnhum.2012.00314
Varese, F., Barkus, E., & Bentall, R. P. (2011). Dissociative and metacognitive factors in hallucination-proneness when controlling for comorbid symptoms. *Cognitive Neuropsychiatry*, *16*(3), 193–217. http://doi.org/10.1080/13546805.2010.495244
Vercammen, A., de Haan, E. H. F., & Aleman, A. (2008). Hearing a voice in the noise: Auditory hallucinations and speech perception. *Psychological Medicine*, *38*(8), 1177–1184. http://doi.org/10.1017/S0033291707002437
Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191–1207. http://doi.org/10.3758/s13428-012-0314-x
Waters, F., Allen, P., Aleman, A., Fernyhough, C., Woodward, T. S., Badcock, J. C., … Laroi, F. (2012). Auditory hallucinations in schizophrenia and nonschizophrenia populations: A review

and integrated model of cognitive mechanisms. *Schizophrenia Bulletin*, *38*(4), 683–692. http://doi.org/10.1093/schbul/sbs045

Waters, F., & Badcock, J. C. (2009). Memory for speech and voice identity in schizophrenia. *The Journal of Nervous and Mental Disease*, *197*(12), 887–891. http://doi.org/10.1097/NMD.0b013e3181c29a76

Watson, L. A., Dritschel, B., Obonsawin, M. C., & Jentzsch, I. (2007). Seeing yourself in a positive light: Brain correlates of the self-positivity bias. *Brain Research*, *1152*(1), 106–110. http://doi.org/10.1016/j.brainres.2007.03.049

Wechsler, D. (1997). *WAIS-III administration and scoring manual*. San Antonio, TX: The Psychological Corporation.

Wilson, M. (1988). MRC psycholinguistic database: Machine-usable dictionary, version 2. 00. *Behavior Research Methods, Instruments, & Computers*, *20*(1), 6–10. http://doi.org/10.3758/BF03202594

Woodruff, P. W., Wright, I. C., Bullmore, E. T., Brammer, M., Howard, R. J., Williams, S. C., … Murray, R. M. (1997). Auditory hallucinations and the temporal cortical response to speech in schizophrenia: A functional magnetic resonance imaging study. *The American Journal of Psychiatry*, *154*(12), 1676–1682. http://doi.org/10.1176/ajp.154.12.1676

Woods, A., Jones, N., Alderson-Day, B., Callard, F., & Fernyhough, C. (2015). Experiences of hearing voices: Analysis of a novel phenomenological survey. *The Lancet Psychiatry*, *2*(4), 323–331. http://doi.org/10.1016/S2215-0366(15)00006-1

Xu, M., Homae, F., Hashimoto, R.-I., & Hagiwara, H. (2013). Acoustic cues for the recognition of self-voice and other-voice. *Frontiers in Psychology*, *4*(735), 1–7. http://doi.org/10.3389/fpsyg.2013.00735

Zhang, Z., Shi, J., Yuan, Y., Hao, G., Yao, Z., & Chen, N. (2008). Relationship of auditory verbal hallucinations with cerebral asymmetry in patients with schizophrenia: An event-related fMRI study. *Journal of Psychiatric Research*, *42*(6), 477–486. http://doi.org/10.1016/j.jpsychires.2007.04.003

Appendix 1

List of words used in the experiment.

NEGATIVE	NEUTRAL	POSITIVE
Abnormal	Actual	Adorable
Afraid	Airy	Alive
Angry	Aloof	Beautiful
Awkward	Ample	Blessed
Bad	Annual	Brave
Beaten	Automatic	Bright
Blind	Average	Brilliant
Bloody	Basic	Calm
Confused	Blank	Careful
Crazy	Blond	Caring
Creepy	Blue	Charming
Cruel	Bold	Clean
Damaged	Brief	Confident
Dead	Broad	Cute
Dirty	Brown	Divine
Dreadful	Casual	Elegant
Dumb	Central	Fabulous
Enraged	Civil	Faithful
Failed	Classic	Famous
Faulty	Close	Fantastic
Fearful	Collected	Free
Foolish	Common	Friendly
Furious	Compact	Funny

Guilty	Constant	Gentle
Helpless	Cubic	Gifted
Horrid	Curly	Glad
Hostile	Daily	Good
Ill	Deep	Gorgeous
Infected	Dry	Gracious
Inferior	Familiar	Grateful
Insane	Flat	Handsome
Jealous	Full	Happy
Lazy	Herbal	Healthy
Lonely	High	Honest
Lost	Informal	Hopeful
Mad	Involved	Incredible
Malign	Large	Inspired
Mean	Lay	Joyful
Messy	Local	Kind
Morbid	Long	Loved
Nasty	Main	Lovely
Nervous	Mild	Loyal
Painful	Mutual	Lucky
Pathetic	Narrow	Magical
Poor	Near	Merry
Punished	Neutral	Nice
Rejected	Open	Perfect
Rude	Overt	Playful
Sad	Plain	Precious
Scabby	Plural	Pretty
Scared	Private	Protected
Selfish	Purple	Proud
Shabby	Quiet	Pure
Shamed	Red	Relaxed
Sick	Regular	Romantic
Sinful	Related	Safe
Sneaky	Round	Satisfied
Stinking	Sharp	Secure
Stupid	Slim	Sexy
Terrible	Small	Slender
Terrified	Square	Special
Tragic	Straight	Splendid
Ugly	Subtle	Strong
Unhappy	Thick	Super
Upset	Tiny	Terrific
Useless	Usual	Thoughtful
Violent	Wet	Truthful
Weak	White	Useful
Wicked	Wild	Wealthy
Wrong	Yellow	Wise