Received: 10 June 2016 | Revised: 30 January 2017 | Accepted: 31 January 2017

DOI: 10.1111/psyp.12880

ORIGINAL ARTICLE

WILEY PSYCHOPHYSIOLOGY SPR

# Emotional self–other voice processing in schizophrenia and its relationship with hallucinations: ERP evidence

**Ana P. Pinheiro[1,2] | Neguine Rezaii[3] | Andréia Rauber[4] | Paul G. Nestor[5] | Kevin M. Spencer[3] | Margaret Niznikiewicz[3]**

[1]Faculty of Psychology, University of Lisbon, Lisbon, Portugal

[2]Neuropsychophysiology Laboratory, CIPsi, School of Psychology, University of Minho, Braga, Portugal

[3]VA Boston Healthcare System, Department of Psychiatry, Harvard Medical School, Boston, Massachusetts

[4]Department of Linguistics, University of Tübingen, Tübingen, Germany

[5]Laboratory of Applied Neuropsychology, College of Liberal Arts, University of Massachusetts, Boston, Massachusetts

**Correspondence**
Ana P. Pinheiro, Neuropsychophysiology Laboratory, CIPsi, School of Psychology, University of Minho, Campus de Gualtar, 4710-957 Braga, Portugal.
Email: ana.pinheiro@psi.uminho.pt

**Funding information**
Science and Technology Foundation (FCT, Portugal) (grants IF/00334/2012, PTDC/PSI-PCL/116626/2010, PTDC/MHC-PCN/0101/2014), FEDER (*European Regional Development Fund*) through the European programs QREN (*National Strategic Reference Framework*), and COMPETE (*Operational Programme 'Thematic Factors of Competitiveness'*) (awarded to A. P. P.)

**Abstract**

Abnormalities in self–other voice processing have been observed in schizophrenia, and may underlie the experience of hallucinations. More recent studies demonstrated that these impairments are enhanced for speech stimuli with negative content. Nonetheless, few studies probed the temporal dynamics of self versus nonself speech processing in schizophrenia and, particularly, the impact of semantic valence on self–other voice discrimination. In the current study, we examined these questions, and additionally probed whether impairments in these processes are associated with the experience of hallucinations. Fifteen schizophrenia patients and 16 healthy controls listened to 420 prerecorded adjectives differing in voice identity (self-generated [SGS] versus nonself speech [NSS]) and semantic valence (neutral, positive, and negative), while EEG data were recorded. The N1, P2, and late positive potential (LPP) ERP components were analyzed. ERP results revealed group differences in the interaction between voice identity and valence in the P2 and LPP components. Specifically, LPP amplitude was reduced in patients compared with healthy subjects for SGS and NSS with negative content. Further, auditory hallucinations severity was significantly predicted by LPP amplitude: the higher the SAPS "voices conversing" score, the larger the difference in LPP amplitude between negative and positive NSS. The absence of group differences in the N1 suggests that self–other voice processing abnormalities in schizophrenia are not primarily driven by disrupted sensory processing of voice acoustic information. The association between LPP amplitude and hallucination severity suggests that auditory hallucinations are associated with enhanced sustained attention to negative cues conveyed by a nonself voice.

**KEYWORDS**

auditory hallucinations, schizophrenia, self, speech, valence, voice identity

## 1 | INTRODUCTION

Abnormalities in self–other voice processing have been reported in psychiatric disorders such as schizophrenia (van der Weiden, Prikken, & van Haren, 2015), and are often enhanced in patients with auditory hallucinations compared to patients without hallucinations and healthy controls (e.g., Allen et al., 2004, 2007; Johns et al., 2001; Pinheiro, Rezaii, Rauber, & Niznikiewicz, 2016; Stephane, Kuskowski, McClannahan, Surerus, & Nelson, 2010).

Auditory hallucinations (the subjective perception of sounds in the absence of an external stimulus; Woodruff, 2004) are experienced by up to 70% of schizophrenia patients, most often as voices (auditory verbal hallucinations, AVH; Nayani & David, 1996). Similar to the experience of listening to an external voice, hallucinated voices convey information

  

about the identity and emotional content of the message, suggesting some overlap of the mechanisms involved in AVH and human voice perception (e.g., Badcock & Chhabra, 2013). However, it is possible that the experience of AVH is associated with specific abnormalities in voice information processing pathways (i.e., identity, speech, affect; Belin, Fecteau, & Bédard, 2004). For example, AVH have been more strongly and consistently associated with alterations in the processing of voice identity and semantic content of speech (see Conde, Gonçalves, & Pinheiro, 2016, for a review).

Despite many attempts to clarify the pathological mechanisms underpinning the experience of "hearing voices" (Badcock & Hugdahl, 2012; David, 2011; Ditman & Kuperberg, 2005; Fernyhough, 2004; Jones, 2010; Modinos et al., 2013; Stephane, Barton, & Boutros, 2001; van Lutterveld, Sommer, & Ford, 2011; Waters et al., 2012; reviewed in Allen, Laroi, McGuire, & Aleman, 2008), AVH remain one of the most intriguing psychopathological symptoms. In recent years, there has been increasing support for the hypothesis that AVH result from a failure in monitoring self-generated speech (e.g., inner speech) resulting in the external misattribution of the self-voice (i.e., a self-produced vocalization is wrongly attributed to an external source: "This is not my voice," e.g., Allen et al., 2007; Johns et al., 2001; Jones, 2010).

Abnormalities in the corollary discharge neural mechanism (that codes for the expected sensation, e.g., the sound of one's own voice during speech production, resulting from the action of contracting the muscles of the vocal tract to produce speech; Ford & Mathalon, 2005) offer a plausible biological explanation for the confusion between self-generated and externally generated auditory feedback ("Is this my voice or yours?") in AVH (e.g., Ford & Mathalon, 2005; Swiney & Sousa, 2014). ERP studies demonstrated that the activity of the auditory cortex is suppressed to the self-voice compared to someone else's voice, as reflected in reduced N1 amplitude to self-generated stimuli (e.g., Heinks-Maldonado, Mathalon, Gray, & Ford, 2005). Effects were also found in the P2 component, even though these have been less consistent; for example, some studies report suppression of P2 amplitude to self-generated sounds (e.g., Knolle, Schröger, & Kotz, 2013), whereas others report no differences between talking and listening to a prerecorded self-voice (Wang et al., 2014). Sensory suppression to self-generated stimuli seems to be impaired in schizophrenia patients. They show reduced N1 suppression in response to real-time feedback of their own voice (e.g., Ford et al., 2001), and these abnormalities are enhanced in patients experiencing AVH compared to patients without AVH and healthy controls, as demonstrated by trait studies[1] (e.g., Heinks-Maldonado et al., 2007).

Studies that probed the neurocognitive mechanisms underlying prerecorded self-generated speech processing in the absence of an active speech production condition revealed that top-down factors, such as the evaluation of the emotional significance of a speech stimulus, may explain impaired self-other voice discrimination and its relationship to AVH (e.g., Daalman, Verkooijen, Derks, Aleman, & Sommer, 2012; Johns et al., 2001; Larøi, Van Der Linden, & Marczewski, 2004; Pinheiro, Rezaii, Rauber, & Niznikiewicz, 2016). Specifically, these studies provided behavioral evidence for a negativity bias in schizophrenia patients experiencing AVH, as reflected in increased error rates in the recognition of self-generated speech (SGS) with negative content. External misattributions leading to recognizing SGS as nonself speech tend to be increased for negative words in schizophrenia patients with AVH relative to healthy subjects or patients in remission (Costafreda, Brébion, Allen, McGuire, & Fu, 2008; Pinheiro, Rezaii, Rauber, & Niznikiewicz, 2016). Moreover, when producing speech and hearing a distorted version of their own voice, patients experiencing AVH, but not healthy controls (HC), misrecognize distorted SGS more often when stimuli are derogatory words (Johns et al., 2001). Abnormal salience detection from speech stimuli may, therefore, contribute to deficits in monitoring the source of voice identity, contributing to internal events being misrecognized as externally generated (Shea et al., 2007).

Besides top-down effects on self–other voice discrimination impairments in schizophrenia, the role of altered bottom-up sensory processes has also been highlighted. For example, Chhabra, Badcock, Maybery, and Leung (2012) demonstrated that schizophrenia patients had difficulties in differentiating between speaker identities due to reduced use of formant dispersion information (i.e., the average frequency difference between formants; Fitch, 1997) but not pitch-based cues. These findings indicate that voice identity cues are processed differently by schizophrenia patients, and that difficulties in discriminating speakers as a function of voice identity may be due to reduced use of resonance-based voice cues.

The studies reviewed above suggest that the processing of self versus nonself voice distinctions does not occur independently of stimulus valence[2] (i.e., emotional vs. neutral semantic content) in schizophrenia. However, the temporal dynamics of self–other speech discrimination in schizophrenia and the role played by semantic valence during distinct stages of self versus nonself voice processing are yet to be determined.

In our previous ERP study with healthy subjects, we probed how listeners integrate *what* is being said with *who* is

[1]In AVH research, state studies compare periods of presence versus absence of AVH (within-subject design), whereas trait studies compare patients who report AVH experience with patients without AVH or healthy controls (between-subjects design; Kühn & Gallinat, 2012).

[2]The term *semantic valence* is used throughout the manuscript to indicate neutral versus emotional semantic content, as opposed to valence associated with prosodic content.

**TABLE 1** Participants' sociodemographic, cognitive, and clinical characteristics

| Variable | Healthy controls Mean (*SD*) | Schizophrenia patients Mean (*SD*) | *t*, *p* |
|---|---|---|---|
| Sociodemographic | | | |
| Age (years) | 48.63 (5.08) | 47.40 (8.48) | 0.492, > .05 |
| Education (years) | 15.31 (1.74) | 13.86 (2.35) | 1.933, >.05 |
| Parental socioeconomic status | 2.38 (1.02) | 2.64 (1.39) | −0.199, > .05 |
| Handedness | 0.77 (0.21) | 0.87 (0.17) | −0.429, > .05 |
| Cognitive | | | |
| Verbal Comprehension Index | 104.75 (17.74) | 94.62 (13.05) | 1.637, > .05 |
| Clinical | | | |
| Illness duration | n.a. | 18.56 (10.63) | n.a. |
| Schizophrenia subtype | n.a. | Paranoid = 4<br>Schizoaffective = 3<br>Unspecified = 8 | n.a. |
| CPZ[a] equivalent | n.a. | 396.91 (327.06) | n.a. |
| PANSS Positive | n.a. | 23 (9.24) | n.a. |
| PANSS Negative | n.a. | 24 (9.32) | n.a. |
| PANSS General | n.a. | 43 (15.79) | n.a. |
| PANSS Total | n.a. | 91 (29.83) | n.a. |
| Global SANS | n.a. | 11 (6.03) | n.a. |
| Global SAPS | n.a. | 11 (3.40) | n.a. |

*Note*. Symptom severity in patients was assessed with the Positive and Negative Syndrome Scale (PANSS), the Scale for the Assessment of Negative Symptoms (SANS), and the Scale for the Assessment of Positive Symptoms (SAPS). n.a. = not applicable.
[a]Chlorpromazine equivalent.

saying it (Pinheiro, Rezaii, Nestor et al., 2016). We examined ERP data from 16 healthy subjects who listened to prerecorded words (self vs. nonself generated) varying in emotional valence (e.g., *brief, pretty, rude*). The participants' task was to decide whether the voice they heard was their own, someone else's voice, or whether they were unsure. Compared with nonself speech (NSS), SGS with neutral valence elicited larger N1 amplitude, SGS with positive valence elicited larger P2 amplitude, whereas both positive and negative SGS elicited larger late positive potential (LPP) amplitude. By showing that self-relevance ("my voice") modulates emotional language comprehension, these findings lend support to previous accounts of a close relationship between processing of self-related information and the processing of emotional valence (e.g., Waters, Allen et al., 2012).

In the current study, we extended our previous work by probing the time course of SGS versus NSS processing in schizophrenia and, specifically, the role of semantic valence in self–other voice discrimination. Additionally, we examined the relationship between these processes and the experience of hallucinations. We hypothesized differences in the processing of voice identity between schizophrenia patients and HC (e.g., Allen et al., 2004, 2007; Johns et al., 2001). Specifically, if the hypothesized impairments are related to altered sensory processing of voice acoustic information, group differences will be observed in early auditory N1 component, irrespective of valence type, reflected in reduced N1 amplitude in the schizophrenia group. If the putative impairments are driven by deficits in the early detection of salience from SGS and NSS stimuli (here, understood as either the emotional relevance of the voice or its self-relevance, i.e., self vs. nonself), then group differences will be observed in the P2, a component that reflects early stimulus categorization, being sensitive to attention (Crowley & Colrain, 2004; Paulmann, Bleichner, & Kotz, 2013) as well as to stimulus saliency (e.g., Pinheiro, Rezaii, Nestor et al., 2016). Specifically, P2 amplitude should be increased in the schizophrenia group reflecting impaired early salience detection from voice stimuli (Pinheiro et al., 2013, 2014). Finally, if the impairments occur mainly at higher-order cognitive processes related to integrating acoustic information with memory-driven representations and elaborative processing of voice

**TABLE 2** Acoustic properties of self-generated vocal stimuli in healthy controls (HC) and schizophrenia patients (SZ), and of nonself vocal stimuli

| Word Valence | | HC Mean (*SD*) | SZ Mean (*SD*) | NS male | NS female |
|---|---|---|---|---|---|
| Neutral words | Duration (ms) | 585.85 (43.86) | 641.78 (79.50) | 583.50 | 654.75 |
| | Mean F0 (Hz) | 136.22 (30.41) | 148.41 (28.26) | 110.24 | 206.58 |
| | Mean intensity (dB) | 73.01 (3.30) | 74.94 (2.75) | 72.66 | 75.86 |
| Positive words | Duration (ms) | 631.94 (58.90) | 683.98 (69.19) | 628.79 | 721.70 |
| | Mean F0 (Hz) | 135.57 (30.68) | 148.25 (28.11) | 101.18 | 211.97 |
| | Mean intensity (dB) | 72.59 (3.38) | 74.38 (2.84) | 71.97 | 75.34 |
| Negative words | Duration (ms) | 622.38 (51.38) | 673.49 (74.63) | 611.65 | 703.02 |
| | Mean F0 (Hz) | 137.36 (30.23) | 147.47 (26.41) | 117.89 | 206.59 |
| | Mean intensity (dB) | 78.18 (20.74) | 74.26 (2.68) | 72.19 | 75.92 |

*Note*. 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 quantization. After the recording session, each word was segmented using Praat software (Boersma & Weenink, 2013). First, voice stimuli were normalized according to peak amplitude by means of a Praat script. Acoustic noise was reduced using a Fourier-based noise reduction algorithm (noise reduction = 14 dB; frequency smoothing = 150 Hz; attack/decay time = 0.15 s) implemented in Audacity 2.0.2 software (http://audacity.sourceforge.net/). NS = nonself; F0 = fundamental frequency.

information, then differences will be observed in later processing stages indexed by the LPP, and reflected in reduced amplitude in the schizophrenia group. Furthermore, considering previous evidence supporting a negativity bias in SGS processing in schizophrenia (Costafreda et al., 2008; Johns et al., 2001) and the role played by emotional processing abnormalities in the pathophysiology of AVH (e.g., Sanjuan et al., 2007), we hypothesized that ERP abnormalities in SGS processing would be valence dependent; that is, they would be enhanced for words with negative valence (e.g., *ugly*) compared with words with neutral (e.g., *round*) or positive (e.g., *beautiful*) valence. Finally, within the patient group, we hypothesized that ERP abnormalities would be associated with the severity of auditory hallucinations, particularly in later ERP components that reflect higher-order evaluation processes. This would dovetail with previous trait studies indicating that abnormalities in voice identity and emotional processing abnormalities are enhanced in schizophrenia patients experiencing AVH (e.g., Costafreda et al., 2008).

## 2 | METHOD

### 2.1 | Participants

Fifteen chronic schizophrenia patients (American Psychiatric Association, 2000) and 16 HC matched for age, sex, handedness, and parental socioeconomic status (Hollingshead, 1975) participated in this study. Subjects had no history of hearing loss. Patients were recruited at the Veterans Affairs Administration Hospital–Brockton, from inpatient and outpatient units. Comparison subjects were recruited from advertisements in local newspapers.

The inclusion criteria were English as first language; right-handedness (Oldfield, 1971); no history of neurological illness; no history of DSM-IV diagnosis of drug or alcohol abuse (APA, 2000); verbal intelligence quotient above 85 (Wechsler, 1997); no hearing, vision, or upper body impairment. For HC, additional inclusion criteria were no history of Axis I or II disorders (First, Gibbon, & Spitzer, 1997; First, Spitzer, Gibbon, & Williams, 2002); no history of Axis I disorder in first- or second-degree family members (Andreasen, Endicott, Spitzer, & Winokur, 1977) (Table 1). Symptom severity in patients was assessed with the Positive and Negative Syndrome Scale (PANSS; Kay, Fiszbein, & Opler, 1987), the Scale for the Assessment of Negative Symptoms (SANS; Andreasen, 1983), and the Scale for the Assessment of Positive Symptoms (SAPS; Andreasen, 1984) (Table 1).

All participants had the procedures fully explained to them and read and signed an informed consent form to confirm their willingness to participate in the study (following Harvard Medical School and Veterans Affairs Boston Healthcare System guidelines).

### 2.2 | Stimuli

Stimuli were adjectives with neutral (70), positive (70), and negative (70) semantic valence. Words were matched for

frequency, concreteness, familiarity, imageability, and number of syllables (online supporting information Figure S1).

## 2.3 | Procedure

Each subject participated in two experimental sessions. The first session involved the recording of the participant's voice. Each participant was asked to read aloud a list of 210 adjectives with neutral or emotional valence (SGS condition). For the NSS condition, the same 210 words were recorded by a male (age = 43) or female (age = 44) native speaker[3] of American English unknown to the participants. The words were spoken with neutral intonation and constant voice intensity (Table 2).

In the ERP experiment, 420 adjectives were presented—210 previously recorded by the participant and 210 previously recorded by an individual unknown to the participant. The identity (self/nonself) and semantic valence (negative/positive/neutral) of speech varied across trials, with 70 words per condition.

The six combinations of speech identity and valence were ordered pseudorandomly and presented in two lists, with the constraint of no more than three consecutive trials of the same condition. Half of the participants received the lists in AB sequence, and half in BA sequence. Stimuli were not repeated as stimulus repetition may attenuate the amplitude of ERP components, primarily early components, due to habituation effects (as demonstrated by studies on affective processing, e.g., Carretié, Hinojosa, & Mercado, 2003; Codispoti, Ferrari, & Bradley, 2007; or by auditory processing studies, e.g., Sambeth, Maes, Quiroga, & Coenen, 2004).

In the ERP session, before each word onset, a fixation cross was presented centrally on the screen for 1,500 ms, and was kept during word presentation to minimize eye movements. After a 1,000-ms interstimulus interval, a question mark signaled 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. 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). The availability of an "unsure" option allowed participants to make a choice between "self" and "other" with some degree of confidence. Buttons of the response pad were also marked with the S (self), O (other), and U (unsure) letters to minimize memory demands. Order of buttons was counterbalanced. Before each experimental block, participants were given a brief training with feedback. The task lasted approximately 50 min (breaks included).

[3]For a male participant, a male control (nonself) voice was used; for a female participant, a female control (nonself) voice was used.

## 2.4 | EEG data acquisition and analysis

The EEG was recorded with a 64-channel BioSemi Active2 system (BioSemi B.V., Amsterdam, Netherlands), and acquired in a continuous mode at a digitization rate of 512 Hz, with a band-pass of 0.01 to 100 Hz.

EEG data were analyzed using BrainVision Analyzer 2 software (Brain Products, Munich, Germany). The EEG channels were referenced offline to the average of the left and right mastoids. EEG data were high-pass filtered with a 0.1 Hz filter. Individual ERP epochs associated with correct responses were created for each stimulus type (SGS neutral, SGS positive, SGS negative, NSS neutral, NSS positive, NSS negative), with −200 ms prestimulus baseline and 1,000 ms poststimulus epoch. The EEG was baseline-corrected using the −200 to 0 ms prestimulus interval. The EEG channels were corrected for vertical and horizontal eye movements using the method of Gratton, Coles, & Donchin (1983). EEG epochs exceeding ± 100 microvolts were not included in individual ERP averages. After artifact rejection, at least 70% of the trials per condition per subject entered the analyses. There were no differences in the number of trials included for each condition for each group ($p > .05$). Three components were selected for the analysis: N1, P2, and LPP. Due to group differences in the latency of the P2, different time windows were selected for each group. Mean amplitude was measured in the 130–210 ms (N1), 215–380 ms (P2 HC), 250–415 ms (P2 patients), and 500–700 ms latency windows (LPP).

## 2.5 | Statistical analyses

### 2.5.1 | ERP data

Repeated measures analyses of variance (ANOVAs) were separately computed for N1, P2, and LPP mean amplitude. Based on a careful inspection of grand-averaged waveforms, a region of interest (ROI) analysis was applied. Our first analysis aimed to test the regional distribution of effects of identity and valence across groups, considering the fact that the N1 and P2 effects have a predominant frontocentral/central distribution (Crowley & Colrain, 2004), whereas the LPP has typically a more central and centroparietal distribution (Bobes, Martín, Olivares, & Valdés-Sosa, 2000). Therefore, identity (2 levels), valence (3 levels), and ROI (frontocentral: FC3/FC1/FCz/FC2/FC4; central: C3/C1/Cz/C2/C4; centroparietal: CP3/CP1/CPz/CP2/CP4) were included as within-subject factors. Our second analysis probed whether there were hemispheric differences across groups, as previous studies have pointed out hemispheric effects related to identity (Rosa, Lassonde, Pinard, Keenan, & Belin, 2008) and/or to affective processing (e.g., Bobes et al., 2000): identity (2 levels), valence (3 levels), and ROI (left lateral: FC5/C5/CP5; left medial 1: FC3/C3/CP3; left medial 2: FC1/C1/

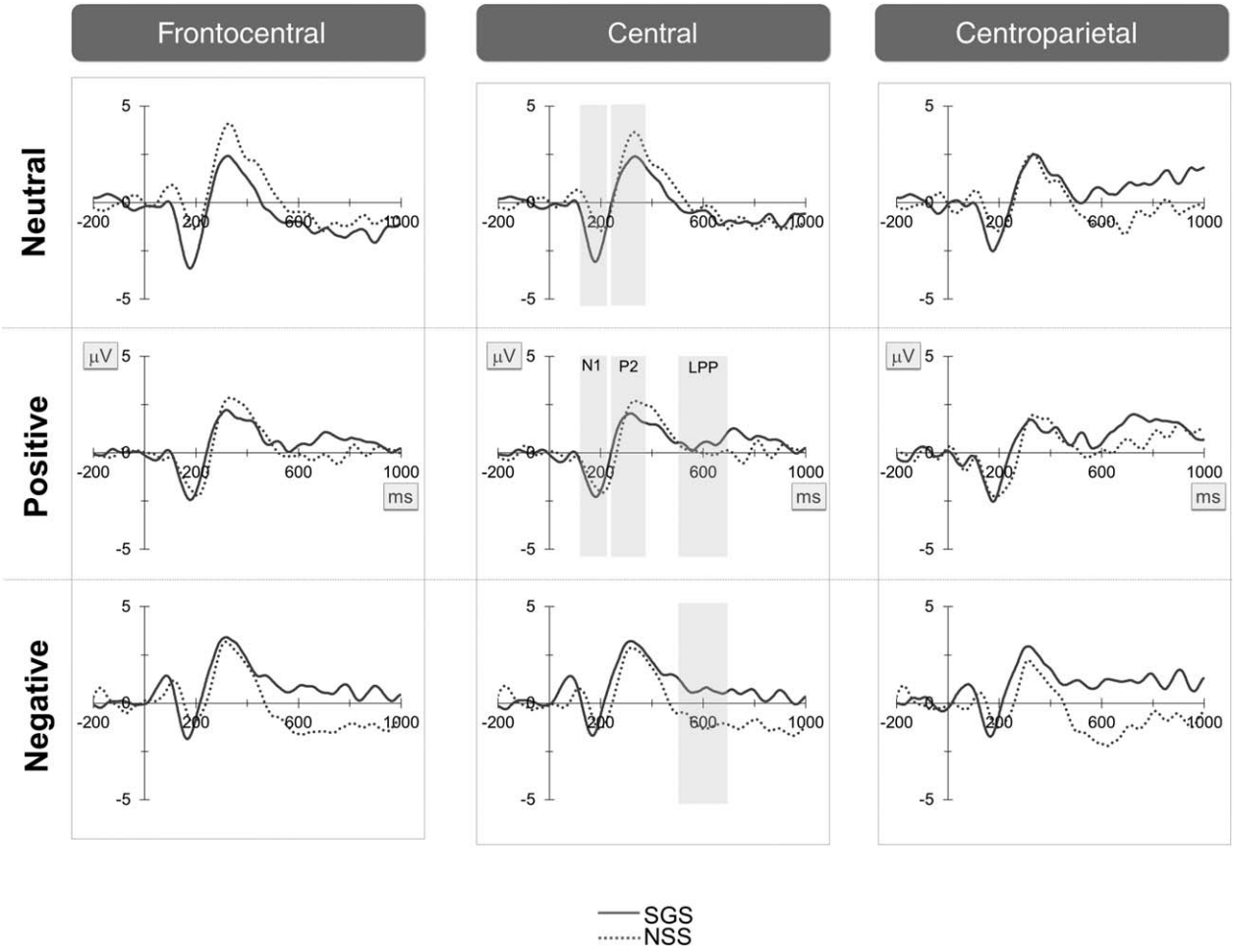


**FIGURE 1** Grand-averaged waveforms showing voice identity (SGS vs. NSS) contrasts for neutral, positive, and negative speech at frontocentral, central, and centroparietal electrode locations in the schizophrenia group. SGS = self-generated speech; NSS = nonself speech; frontocentral = average of amplitude at electrodes F3/F1/Fz/F2/F4; central = average of amplitude at electrodes C3/C1/Cz/C2/C4; centroparietal = average of amplitude at electrodes CP3/CP1/CPz/CP2/CP4

CP1; right lateral: FC6/C6/CP6; right medial 1: FC4/C4/CP4; right medial 2: FC2/C2/CP2) were tested.

Analyses were corrected for nonsphericity using the Greenhouse-Geisser method (the original *df* is reported). Effect sizes for significant effects are reported using the partial η squared method ($\eta_p^2$).

### 2.5.2 | Behavioral data

The effects of identity and valence on voice recognition accuracy and number of unsure responses were tested for each response type separately, using repeated measures ANOVA with identity and valence as within-subject factors, and group as between-subjects factor.

### 2.5.3 | Relationship between ERPs and hallucinations

A generalized linear model with Poisson distribution for significant ERP effects tested the predictive value of ERP amplitude on clinical measures of hallucinations (zero-inflation model).

## 3 | RESULTS

### 3.1 | ERP

Figure 1 presents grand-averaged waveforms in schizophrenia patients (data from healthy controls obtained with the same stimuli are available in Pinheiro, Rezaii, Nestor et al., 2016). As the analysis was focused on group differences, we report significant group effects or interactions involving the group factor (Figure 2, Table 3). In the case of a significant interaction with group factor, our first analysis focused on following up the interaction to probe quantitative group differences in ERP amplitude (between-groups effects). Second, since the emphasis in this study was also on the pattern of condition differences within each group, we followed up significant interactions with the group factor by running within-

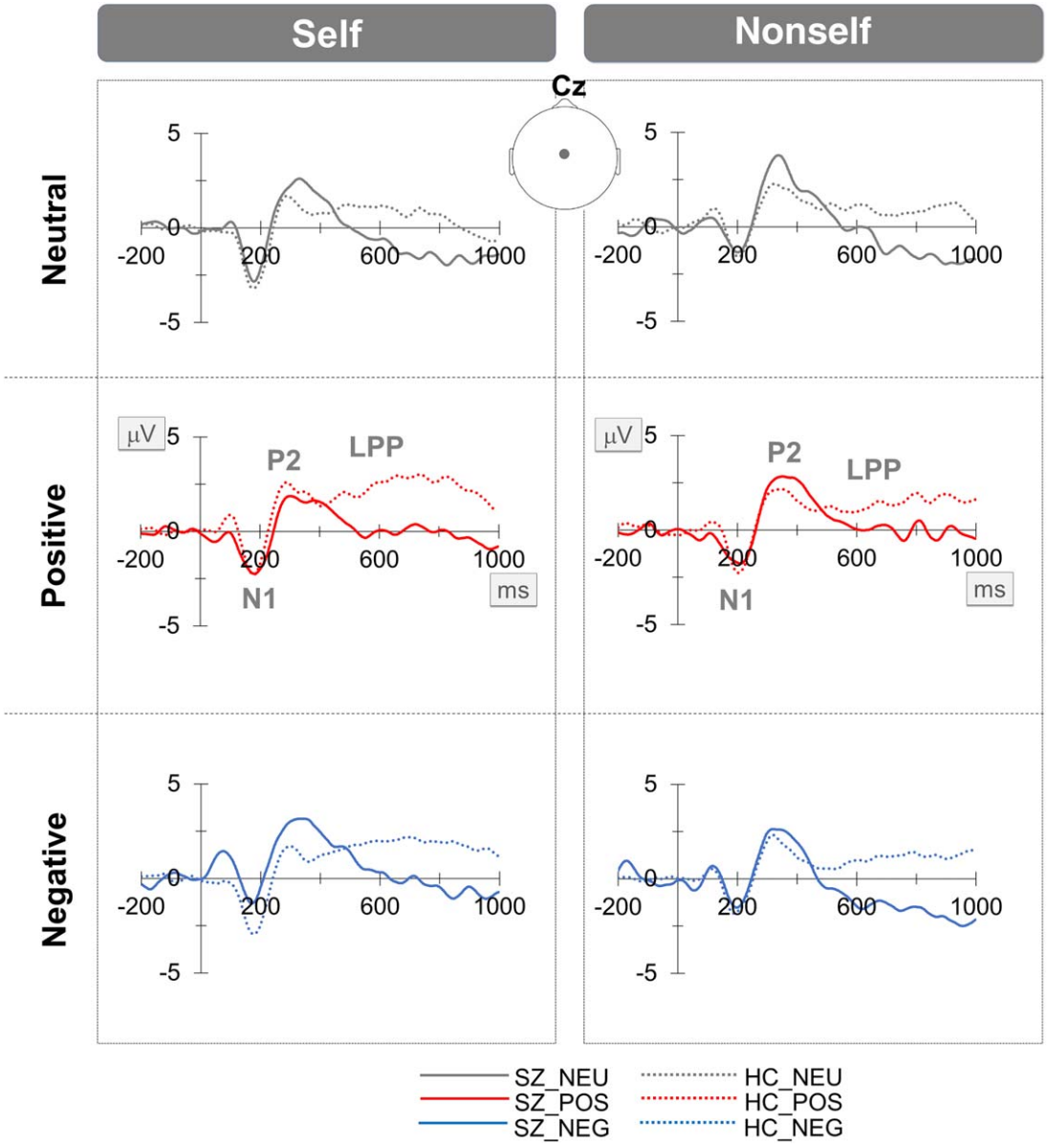


**FIGURE 2** Grand-averaged waveforms showing group contrasts for self- and nonself speech with neutral, positive, and negative valence at the CZ electrode. SZ = schizophrenia; HC = healthy controls; NEU = neutral; POS = positive; NEG = negative

**TABLE 3** Mean amplitude of N1, P2, and LPP components for each experimental condition in healthy controls and schizophrenia patients

| | | SGS | | | NSS | | |
|---|---|---|---|---|---|---|---|
| | | **Neutral** | **Positive** | **Negative** | **Neutral** | **Positive** | **Negative** |
| N100 | HC | −2.774 (.615) | −1.617 (.637) | −2.365 (.652) | −1.307 (.532) | −2.103 (.490) | −1.907 (.525) |
| | SZ | −2.491 (.635) | −1.887 (.658) | −0.795 (.674) | −1.240 (.549) | −2.051 (.506) | −0.917 (.542) |
| P200 | HC | 1.338 (.678) | 2.341 (.824) | 1.517 (.829) | 1.431 (.722) | 1.003 (.887) | 1.039 (.703) |
| | SZ | 1.485 (.700) | 1.363 (.851) | 2.574 (.857) | 2.246 (.746) | 1.086 (.916) | 1.839 (.355) |
| LPP | HC | 1.264 (1.112) | 2.916 (1.159) | 2.401 (1.173) | 1.107 (1.074) | 1.111 (1.345) | .854 (1.131) |
| | SZ | −1.074 (1.149) | 0.186 (1.197) | 0.261 (1.212) | −1.034 (1.109) | −.459 (1.389) | −1.802 (1.168) |

*Note*. Standard errors are provided in parentheses. HC = healthy controls; SZ = schizophrenia patients.

group ANOVAs testing condition effects in patients and controls separately (within-group effects).

## 3.2 | N1

No significant group effects were observed ($p > .05$).

## 3.3 | P2

A main effect of ROI, $F(2, 28) = 21.421$, $p < .001$, $\eta_p^2 = .605$, revealed larger P2 amplitude in both frontocentral and central regions than in the centroparietal region (frontocentral > centroparietal: $p < .001$; central > centroparietal: $p < .001$). Importantly, a significant Group × Identity × Valence × ROI interaction: $F(4, 26) = 3.157$, $p = .031$, $\eta_p^2 = .327$, was observed.

### 3.3.1 | Between-groups effects

The interaction was followed up by computing separate ANOVAs that tested the effects of voice identity for each valence type (i.e., neutral, positive, and negative speech). All other factors were kept the same. Group contrasts were not significant ($p > .05$).

### 3.3.2 | Within-group effects

We also examined the effects of voice identity in each group, considering each valence type separately. In HC, P2 for positive speech was enhanced in the case of SGS relative to NSS (identity effect: $F(1, 15) = 4.495$, $p = .05$, $\eta_p^2 = .231$), while there were no significant identity differences in the case of neutral or negative speech ($p > .05$). In schizophrenia, voice identity differences were found for neutral speech only, and depended on the region (Identity × ROI interaction: $F(2, 13) = 5.310$, $p = .021$, $\eta_p^2 = .450$): P2 was enhanced for neutral NSS relative to neutral SGS at frontocentral electrode sites ($p = .025$).

### 3.3.3 | Hemispheric effects

The hemispheric distribution of the P2 effects did not differ across groups ($p > .05$).

## 3.4 | LPP

The repeated measures ANOVA yielded a four-way interaction between group, identity, valence, and ROI: $F(6, 24) = 4.381$, $p = .037$, $\eta_p^2 = .946$.

### 3.4.1 | Between-groups effects

We followed up the interaction by running separate ANOVAs for each valence type and ROI. LPP was decreased for negative SGS and negative NSS in schizophrenia relative to HC at centroparietal electrode sites (group: $F(1, 29) = 4.278$, $p = .048$, $\eta_p^2 = .129$).

### 3.4.2 | Within-group effects

We probed the interactions between identity, valence, and ROI in each group separately. In HC, LPP was increased for positive SGS relative to positive NSS (identity effect: $F(1, 15) = 16.274$, $p = .001$, $\eta_p^2 = .520$), and for negative SGS relative to negative NSS (identity effect: $F(1, 15) = 7.190$, $p = .017$, $\eta_p^2 = .324$). In schizophrenia, LPP amplitude was similar for neutral SGS and NSS, and for positive SGS and NSS ($p > .05$); however, LPP was larger for negative SGS relative to negative NSS (identity effect: $F(1, 14) = 8.455$, $p = .011$, $\eta_p^2 = .377$).

### 3.4.3 | Hemispheric effects

The hemispheric distribution of the LPP effects did not differ across groups ($p > .05$).

## 3.5 | Behavioral data

No significant group effect was observed ($p > .05$). However, there was a trend for a Group × Valence interaction, $F(2, 56) = 2.876$, $p = .065$, $\eta_p^2 = .093$. In order to clarify this trend, we explored within-group effects. No significant effect of identity or valence was found in HC. However, a significant valence effect characterized the performance of the schizophrenia group, $F(2, 26) = 4.887$, $p = .016$, $\eta_p^2 = .273$: patients were less accurate in the recognition of negative relative to positive speech ($p = .049$; Figure 3, Table 4).

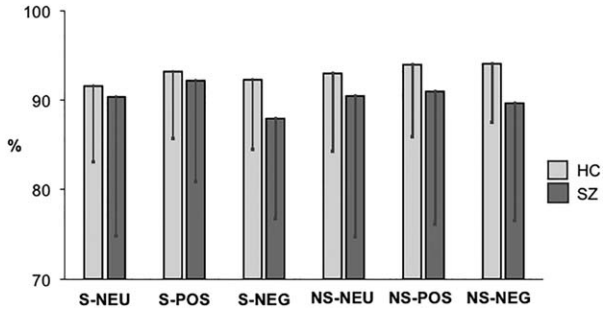


**FIGURE 3** Accuracy in the recognition of self- versus nonself speech in healthy controls (HC) and schizophrenia patients (SZ). Standard deviations are represented in the figure by the error bars attached to each column

**TABLE 4** Percentage of hits and unsure responses as a function of voice identity and semantic valence in healthy controls and schizophrenia patients

| | | SGS | | | NSS | | |
|---|---|---|---|---|---|---|---|
| | | Neutral | Positive | Negative | Neutral | Positive | Negative |
| Hits | HC | 91.61 (8.52) | 93.21 (7.48) | 92.32 (7.84) | 93.04 (8.70) | 94.02 (8.12) | 94.11 (6.55) |
| | SZ | 90.38 (15.49) | 92.19 (11.26) | 88.00 (11.21) | 90.48 (15.70) | 90.95 (14.79) | 89.71 (13.12) |
| | Total | 91.01 (12.20) | 92.72 (9.35) | 90.23 (9.70) | 91.80 (12.43) | 92.53 (11.73) | 91.98 (10.33) |
| Unsure responses | HC | 1.88 (2.26) | 1.52 (1.98) | 1.70 (2.67) | 2.05 (3.00) | 0.98 (1.25) | 1.07 (1.52) |
| | SZ | 4.57 (6.92) | 2.95 (4.02) | 4.38 (5.14) | 2.76 (5.85) | 2.86 (5.89) | 2.95 (6.21) |
| | Total | 3.18 (5.17) | 2.21 (3.17) | 3.00 (4.21) | 2.40 (4.54) | 1.89 (4.23) | 1.98 (4.48) |

*Note*. Mean and standard deviation (in parentheses) values are provided. HC = healthy controls; SZ = schizophrenia patients.

No significant group, voice identity, or condition effects were observed when analyzing the number of unsure responses ($p > .05$).

### 3.6 | Relationship between ERP and clinical data

The regression analysis showed that the LPP amplitude for NSS with negative content only was significantly predicted by PANSS hallucination severity ($p = .044$): the more severe the hallucination score, the more positive the amplitude (see Table 5).

In order to better explore the association between auditory hallucinations and negative speech processing, we computed a difference score between LPP amplitude for negative and positive NSS, as well as a difference score between LPP amplitude for negative NSS and negative SGS. We observed that the LPP negative-positive difference score was significantly predicted by the SAPS voices conversing score: the more severe the auditory hallucination score, the larger the difference (i.e., the more negative the difference outcome; see Table 5).

P2 amplitude did not predict PANSS hallucination severity ($p > .05$).

## 4 | DISCUSSION

This study examined the interactions between voice identity and semantic valence during speech perception in schizophrenia. Corroborating our hypotheses, we found group differences in the processing of SGS and NSS that depended on semantic valence. These differences were observed after 200 ms after speech onset.

### 4.1 | The time course of self–other voice processing and the effects of valence: Group differences

In HC, the P2 response was enhanced for positive SGS compared to positive NSS, whereas in patients enhanced P2

**TABLE 5** Association between LPP and hallucinations

| | Correlation | Regression model | | | |
|---|---|---|---|---|---|
| | (*r*) | Estimate | Standard error | *z* value | Probability (> \|z\|) |
| PANSS hallucinations | | | | | |
| Intercept | n.a. | 0.166 | 0.402 | 0.413 | 0.680 |
| Negative NSS | .41^ | 0.229 | 0.114 | 2.014 | 0.044* |
| SAPS "voices conversing" | | | | | |
| Intercept | n.a. | 0.497 | 0.236 | 2.105 | 0.0353 |
| Difference Negative NSS-Positive NSS | −.39^ | −0.131 | 0.062 | −2.107 | 0.035* |

*Note*. n.a. = not applicable.

*Significant ($p < .05$), after using the Holm-Bonferroni method.

^Marginally significant ($p < .01$), after using the Holm-Bonferroni method.

amplitude was observed for neutral NSS compared to neutral SGS. The literature on the auditory P2 is sparse, and the functional significance of this component remains to be clarified (for a review, see Crowley & Colrain, 2004). However, the existing evidence supports the role of P2 in early and rudimentary stimulus classification and categorization, and has pointed out the modulatory effects of attention on its amplitude (Crowley & Colrain, 2004). There is also some evidence that the P2 is sensitive to stimulus saliency, such as its emotional relevance (e.g., Paulmann & Kotz, 2008; Pinheiro et al., 2013, 2014) or self-relevance/ownership (e.g., whether a sound was produced by oneself or not; Knolle et al., 2013; Timm, Schönwiesner, Schröger, & SanMiguel et al., 2016). These studies indicate that the P2 amplitude reflects attentional orienting that is boosted (increased P2) by salient (emotional/self-relevant) stimuli. Based on this evidence, a putative explanation for the current P2 findings is that they reflect early attentional selection of stimulus features that are processed prior to full access to word meaning, thus allowing a rapid detection of salient events (e.g., Kanske, Plitschka, & Kotz, 2011). In the case of HC, positive words (e.g., *beautiful*, *adorable*) spoken in one's own voice boosted the attention capture effect. Along the same lines, previous studies reported enhanced attention for emotional (positive) compared to neutral verbal stimuli (e.g., enhanced P2: Kanske & Kotz, 2007; Kissler, Assadollahi, & Herbert, 2006), and for self-voice compared to nonself voice stimuli (enhanced P3: Conde, Gonçalves, & Pinheiro, 2015). The P2 finding in HC may additionally suggest that the abstraction of voice identity features for early stimulus discrimination (self vs. nonself) was facilitated in the case of positive speech, which may reflect additive effects of saliency indexed by self-relevance ("my voice") and valence. In contrast, patients showed increased P2 amplitude for NSS, and the differentiation between SGS and NSS occurred only for speech stimuli with neutral content. This finding suggests that attention was more strongly oriented to nonemotional and nonself vocal stimuli at an early processing stage. This finding is consistent with accounts of disrupted salience detection in early stages of speech processing in schizophrenia (e.g., Pinheiro et al., 2013, 2014). Further, it demonstrates that the abstraction of voice identity features that is critical for determining whether a sound was produced by oneself ("This is my voice") or not ("This voice is not mine") differs from healthy controls after 200-ms poststimulus onset.

Interactive effects of voice identity and speech valence were observed in a later stage indexed by the LPP. These effects were manifested both as between- and within-group differences. LPP amplitude was reduced in patients compared with HC for negative SGS and NSS. Furthermore, while HC showed increased amplitude for SGS relative to NSS with both positive and negative valence, patients showed increased amplitude for SGS compared to NSS with negative valence only. The LPP component has been consistently reported in emotion and self-related processing research, and is considered a neurophysiological signature of enhanced and sustained attention to salient (emotional or self-relevant) stimuli (e.g., Tacikowski & Nowicka, 2010). While in controls we replicated a consistent finding in emotion research (i.e., increased LPP for emotionally salient stimuli, e.g., Kissler et al., 2006), the patient group results suggest that sustained attention is specifically increased for negative SGS, that is in a self-relevant ("my voice") context. This finding provides additional support to the negativity bias found in previous studies (e.g., Costafreda et al., 2008; Pinheiro et al., 2013). It is plausible that for patients these stimuli represent more salient stimuli that attract more attention and require increased elaborative processing. In good accordance with this hypothesis, fMRI studies demonstrated enhanced activation of limbic (e.g., the amygdala) and frontal brain regions (e.g., orbitofrontal cortex) in schizophrenia patients with persistent hallucinations in response to emotional words that were the most frequent words heard during the hallucinatory states (e.g., imperative words of negative content, insults; De La Iglesia-Vaya et al., 2014; Escartí et al., 2010; Sanjuan et al., 2007).

### 4.2 | The time course of self–other voice processing and the effects of valence: Group similarities

Contrary to previous accounts of disrupted bottom-up processing of voice information (e.g., Chhabra et al., 2012; Pinheiro et al., 2013, 2014), the sensory processing of speech occurred similarly in schizophrenia patients and healthy subjects regardless of voice identity. In both groups, N1 was enhanced for SGS relative to NSS. Since the auditory N1 component is regarded as a neurophysiological signature of perceptual processing and automatic attention allocation (Rosburg, Boutros, & Ford, 2008), this finding indicates that both groups processed SGS and NSS stimuli similarly at an early processing stage. Differences in task requirements and study measures may explain the apparent discrepancy between the current work and previous studies (e.g., Chhabra et al., 2012). The ERP results reported in the earlier Niznikiewicz et al. study (1997) further underscore the critical role of a study task in the pattern of observed results. In this study, a lack of N1 abnormalities in the patient group in a task of auditory sentence comprehension coexisted with abnormalities in the N400 component (more negative amplitude to both correct and incorrect sentence endings). The current results are also different from those reported in our previous studies of prosody processing (Pinheiro et al., 2013, 2014) in which the study task emphasis was on the

processing of acoustic properties of speech as they related to emotional speech valence. As the current patients' sample does not fully overlap with the samples we tested in our prior studies, the hypothesis that differences in clinical status accounted for the inconsistent findings should not be ruled out. Indeed, the presence of enduring negative symptoms was found to play a major role in the N1 amplitude reduction in schizophrenia patients compared to controls (Mucci et al., 2007).

Furthermore, as we did not control for nicotine use (which tends to be increased in schizophrenia patients compared to healthy control subjects; McCreadie, 2002), we cannot rule out potential nicotinic effects on the N1 amplitude, as nicotine was found to modulate early auditory potentials in the N1 latency range (Baldeweg, Wong, & Stephan, 2006).

### 4.3 | Self–other voice discrimination accuracy and the effects of valence: Group differences

In spite of within- and between-group ERP differences in the processing of self versus nonself speech processing, patients and HC recognized both types of stimuli with equally high accuracy. In both groups, irrespective of identity and valence, the accuracy rate was over 90%, agreeing with more recent studies (Hughes & Nicholson, 2010; Rosa et al., 2008). Nonetheless, the analysis of the qualitative recognition profile in each group revealed valence-dependent effects in schizophrenia but not in the control group: patients were less accurate in recognizing negative SGS as *self* and negative NSS as *other*. This finding provided further support to the negativity bias reported in previous schizophrenia studies (e.g., Costafreda et al., 2008; Johns et al., 2001; Pinheiro et al., 2013).

The effects of task length on performance and attention should be considered. Even though the duration of the task was relatively long (approximately 50 min), the experimental session was broken by several short breaks (a break of approximately 30 s occurred after every 20 trials). Critically, as group differences were specifically observed in the negative NSS condition, it is less likely that task duration had a negative impact on attention. If that were the case, generalized impairments in recognition accuracy would have been observed in the schizophrenia group.

### 4.4 | Self–other negative voice processing and its relationship with hallucinations

The regression analysis showed that the LPP amplitude for negative NSS was associated with hallucination severity. We note, though, that the association is not specific of auditory hallucinations, as the PANSS hallucinations subscale also taps into visual, olfactory, or somatic hallucinations. Despite this limitation, the specific relationship between hallucinations and abnormal negative speech processing is consistent with a negativity bias found in previous self-monitoring studies. In these studies, AVH patients showed increased external misattribution errors for negative affective stimuli (words such as *abnormal* or *monstrous*) when compared to both healthy subjects and patients in remission (Allen et al., 2004; Costafreda et al., 2008; Johns et al., 2001; Pinheiro, Rezaii, Rauber, & Niznikiewicz, 2016).

The additional association between an auditory hallucinations scale (SAPS voices conversing) and the LPP positive-negative NSS difference suggests that emotional self–other voice processing abnormalities may represent a specific feature of patients experiencing auditory verbal hallucinations compared to patients experiencing other types of hallucinations. This association showed that the more severe the hallucination severity (voices conversing), the larger the difference. This seems to suggest that, as voices are more often described as possessing a negative or frightening tone, and thus may cause significant emotional distress (e.g., Nayani & David, 1996), attentional resources may become more tuned to negative information spoken by others. Consequently, negative NSS may be a specific target of increased elaborative processing and sustained attention. Partial support for this hypothesis comes from previous reports of attentional biases toward negative stimuli in schizophrenia and of their contribution to the experience of hallucinations (e.g., Costafreda et al., 2008; Johns et al., 2001). We note, however, that a regression analysis is not sufficient to establish causality.

### 4.5 | Limitations and future directions

Considering the modest sample size in the current study, these findings should be interpreted with caution. Moreover, as abnormal LPP to negative NSS was associated with the PANSS hallucination scale, which does not tap exclusively into auditory hallucinations but also includes visual hallucinations, future studies should probe the putative relationship between altered negative speech processing and auditory hallucinations specifically in a larger patients' sample and using more sensitive measures of hallucinations in the auditory modality. A more detailed assessment of the nature of AVH might capture the complexity of this phenomenon in a more appropriate and comprehensive manner.

### 4.6 | Conclusions

This study demonstrated that the interactive effects of voice identity and semantic valence differed for the P2 and LPP components in schizophrenia and healthy controls. A negativity bias in schizophrenia was reflected in more positive LPP amplitude for nonself relative to self-speech with negative content only. This pattern was confirmed by the behavioral analysis that indicated decreased accuracy when

recognizing the identity of negative spoken words. The absence of group differences in the N1 component suggests that self–other voice identity processing abnormalities are not primarily driven by disrupted sensory processing of voice acoustic information.

Together, the ERP and behavioral data provide support for the contribution of higher-order top-down processes (e.g., evaluation of speech emotional meaning) to abnormal self-other voice processing in schizophrenia. The association between LPP amplitude and hallucination severity suggests that enhanced sustained attention to negative cues conveyed by a non-self voice may play a role in the hallucinatory experience.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge all the participants who collaborated in the study, as well as Dr. Brett Clementz and two anonymous reviewers for their valuable comments to previous versions of the manuscript and suggestions. This work was supported by grants IF/00334/2012, PTDC/PSI-PCL/116626/2010, and PTDC/MHC-PCN/0101/2014 funded by Science and Technology Foundation (FCT, Portugal) and FEDER (European Regional Development Fund) through the European programs QREN (National Strategic Reference Framework), and COMPETE (Operational Program 'Thematic Factors of Competitiveness'), awarded to A. P. P.

## REFERENCES

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. *British Journal of Psychiatry*, *190*, 162–169. Doi: 10.1192/bjp.bp.106.025700

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. Doi: 10.1016/j.schres.2003.09.008

Allen, P., Laroi, F., McGuire, P. K., & Aleman, A. (2008). The hallucinating brain: A review of structural and functional neuroimaging studies of hallucinations. *Neuroscience and Biobehavioral Reviews*, *32*(1), 175–191. Doi: 10.1016/j.neubiorev.2007.07.012

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: American Psychiatric Association Press. Doi: 10.1016/B978-1-4377-2242-0.00016-X

Andreasen, N. C. (1984). *Scale for the Assessment of Positive Symptoms*. Iowa City, IA: University of Iowa.

Andreasen, N. C. (1983). *Scale for the Assessment of Negative Symptoms*. Iowa City, IA: University of Iowa.

Andreasen, N. C., Endicott, J., Spitzer, R. L., & Winokur, G. (1977). The family history method using diagnostic criteria. Reliability and validity. *Archives of General Psychiatry*, *34*(10), 1229–35. Doi: 10.1001/archpsyc.1977.01770220111013

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. Doi: 10.3389/fnhum.2013.00114

Badcock, J. C., & Hugdahl, K. (2012). Cognitive mechanisms of auditory verbal hallucinations in psychotic and non-psychotic groups. *Neuroscience and Biobehavioral Reviews*, *36*(1), 431–438. Doi: 10.1016/j.neubiorev.2011.07.010

Baldeweg, T., Wong, D., & Stephan, K. E. (2006). Nicotinic modulation of human auditory sensory memory: Evidence from mismatch negativity potentials. *International Journal of Psychophysiology*, *59*(1), 49–58. Doi: 10.1016/j.ijpsycho.2005.07.014

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

Bobes, M. A., Martín, M., Olivares, E., & Valdés-Sosa, M. (2000). Different scalp topography of brain potentials related to expression and identity matching of faces. *Cognitive Brain Research*, *9*(3), 249–260. Doi: 10.1016/S0926-6410(00)00003-3

Boersma, P., & Weenink, D. (2013). Praat: Doing phonetics by computer [Computer program]. Version 5.3.53. Retrieved from http://www.praat.org/

Carretié, L., Hinojosa, J. A., & Mercado, F. (2003). Cerebral patterns of attentional habituation to emotional visual stimuli. *Psychophysiology*, *40*(3), 381–388. Doi: 10.1111/1469-8986.00041

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

Codispoti, M., Ferrari, V., & Bradley, M. M. (2007). Repetition and event-related potentials: Distinguishing early and late processes in affective picture perception. *Journal of Cognitive Neuroscience*, *19*(4), 577–586. Doi: 10.1162/jocn.2007.19.4.577

Conde, T., Gonçalves, O. F., & Pinheiro, A. P. (2015). Paying attention to my voice or yours: An ERP study with words. *Biological Psychology*, *111*, 40–52. Doi: 10.1016/j.biopsycho.2015.07.014

Conde, T., Gonçalves, Ó. 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. Doi: 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. Doi: 10.1017/S0033291708003243

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: Age, sleep and modality. *Clinical Neurophysiology*, *115*(4), 732–744. Doi: 10.1016/j.clinph.2003.11.021

Daalman, K., Verkooijen, S., Derks, E. M., Aleman, A., & Sommer, I. E. C. (2012). The influence of semantic top-down processing in auditory verbal hallucinations. *Schizophrenia Research*, *139*(1–3), 82–86. Doi: 10.1016/j.schres.2012.06.005

David, A. S. (2011). The cognitive neuropsychiatry of auditory verbal hallucinations: An overview. *Cognitive Neuropsychiatry*, *9*(1–2), 107–123. Doi: 10.1080/13546800344000183

De La Iglesia-Vaya, M., Escartí, M. J., Molina-Mateo, J., Martí-Bonmatí, L., Gadea, M., Castellanos, F. X., ... Sanjuan, J. (2014). Abnormal synchrony and effective connectivity in patients with

schizophrenia and auditory hallucinations. *NeuroImage: Clinical*, *6*, 171–179. Doi: 10.1016/j.nicl.2014.08.027

Ditman, T., & Kuperberg, G. R. (2005). A source-monitoring account of auditory verbal hallucinations in patients with schizophrenia. *Harvard Review of Psychiatry*, *13*(5), 280–299. Doi: 10.1080/10673220500326391

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. Doi: 10.1016/j.schres.2009.12.028

Fernyhough, C. (2004). Alien voices and inner dialogue: Towards a developmental account of auditory verbal hallucinations. *New Ideas in Psychology*, *22*(1), 49–68. Doi: 10.1016/j.newideapsych.2004.09.001

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.

Fitch, W. T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *Journal of the Acoustical Society of America*, *102*, 1213–1222.

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. Doi: 10.1016/j.ijpsycho.2005.01.014

Ford, J. M., Mathalon, D. H., Heinks, T., Kalba, S., Faustman, W. O., & Roth, W. T. (2001). Neurophysiological evidence of corollary discharge dysfunction in schizophrenia. *American Journal of Psychiatry*, *158*(12), 2069–2071. Doi: 10.1176/appi.ajp.158.12.2069

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*(4), 468–484. Doi: 10.1016/0013-4694(83)90135-9

Heinks-Maldonado, T. H., Mathalon, D. H., Gray, M., & Ford, J. M. (2005). Fine-tuning of auditory cortex during speech production. *Psychophysiology*, *42*(2), 180–90. Doi: 10.1111/j.1469-8986.2005.00272.x

Heinks-Maldonado, T. H., Mathalon, D. H., Houde, J. F., Gray, M., Faustman, W. O., & Ford, J. M. (2007). Relationship of imprecise corollary discharge in schizophrenia to auditory hallucinations. *Archives of General Psychiatry*, *64*(3), 286–296. Doi: 10.1001/archpsyc.64.3.286

Hollingshead, A. (1975). Four factor index of social status. *Yale Journal of Sociology*, *8*, 21–52.

Hughes, S. M., & Nicholson, S. E. (2010). The processing of auditory and visual recognition of self-stimuli. *Consciousness and Cognition*, *19*(4), 1124–1134. Doi: 10.1016/j.concog.2010.03.001

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. Doi: 10.1017/S0033291701003774

Jones, S. R. (2010). Do we need multiple models of auditory verbal hallucinations? Examining the phenomenological fit of cognitive and neurological models. *Schizophrenia Bulletin*, *36*(3), 566–575. Doi: 10.1093/schbul/sbn129

Kanske, P., & Kotz, S. A. (2007). Concreteness in emotional words: ERP evidence from a hemifield study. *Brain Research*, *1148*, 138–148. Doi: 10.1016/j.brainres.2007.02.044

Kanske, P., Plitschka, J., & Kotz, S. A. (2011). Attentional orienting towards emotion: P2 and N400 ERP effects. *Neuropsychologia*, *49*(11), 3121–3129. Doi: 10.1016/j.neuropsychologia.2011.07.022

Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The Positive and Negative Syndrome Scale (PANSS) for Schizophrenia. *Schizophrenia Bulletin*, *13*(2), 261–276. Doi: 10.1093/schbul/13.2.261

Kissler, J., Assadollahi, R., & Herbert, C. (2006). Emotional and semantic networks in visual word processing: Insights from ERP studies. *Progress in Brain Research*, *156*(6), 147–183. Doi: 10.1016/S0079-6123(06)56008-X

Knolle, F., Schröger, E., & Kotz, S. A. (2013). Cerebellar contribution to the prediction of self-initiated sounds. *Cortex*, *49*(9), 2449–2461. Doi: 10.1016/j.cortex.2012.12.012

Kühn, S., & Gallinat, J. (2012). Quantitative meta-analysis on state and trait aspects of auditory verbal hallucinations in schizophrenia. *Schizophrenia Bulletin*, *38*(4), 779–786. Doi: 10.1093/schbul/sbq152

Larøi, F., Van Der Linden, M., & Marczewski, P. (2004). The effects of emotional salience, cognitive effort and meta-cognitive beliefs on a reality monitoring task in hallucination-prone subjects. *British Journal of Clinical Psychology*, *43*, 221–233. Doi: 10.1348/0144665031752970

McCreadie, R. G. (2002). Use of drugs, alcohol and tobacco by people with schizophrenia: Case-control study. *British Journal of Psychiatry*, *181*(4), 321–325. Doi: 10.1192/bjp.181.4.321

Modinos, G., Costafreda, S. G., Van Tol, M. J., McGuire, P. K., Aleman, A., & Allen, P. (2013). Neuroanatomy of auditory verbal hallucinations in schizophrenia: A quantitative meta-analysis of voxel-based morphometry studies. *Cortex*, *49*(4), 1046–1055. Doi: 10.1016/j.cortex.2012.01.009

Mucci, A., Galderisi, S., Kirkpatrick, B., Bucci, P., Volpe, U., Merlotti, E., … Maj, M. (2007). Double dissociation of N1 and P3 abnormalities in deficit and nondeficit schizophrenia. *Schizophrenia Research*, *92*(1–3), 252–261. Doi: 10.1016/j.schres.2007.01.026

Nayani, T. H., & David, A. S. (1996). The auditory hallucination: A phenomenological survey. *Psychological Medicine*, *26*(1), 177–189. Doi: 10.1017/S003329170003381X

Niznikiewicz, M. A., O'Donnell, B. F., Nestor, P. G., Smith, L., Law, S., Karapelou, M., … McCarley, R. W. (1997). ERP assessment of visual and auditory language processing in schizophrenia. *Journal of Abnormal Psychology*, *106*(1), 85–94. Doi: 10.1037/0021-843X.106.1.85

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia*, *9*(1), 97–113. Doi: 10.1016/0028-3932(71)90067-4

Paulmann, S., Bleichner, M., & Kotz, S. A. (2013). Valence, arousal, and task effects in emotional prosody processing. *Frontiers in Psychology*, *4*, 345. Doi: 10.3389/fpsyg.2013.00345

Paulmann, S., & Kotz, S. A. (2008). Early emotional prosody perception based on different speaker voices. *NeuroReport*, *19*(2), 209–213. Doi: 10.1097/WNR.0b013e3282f454db

Pinheiro, A. P., del Re, E., Mezin, J., Nestor, P. G., Rauber, A., McCarley, R. W., … Niznikiewicz, M. A. (2013). Sensory-based and higher-order operations contribute to abnormal emotional prosody processing in schizophrenia: An electrophysiological investigation. *Psychological Medicine*, *43*(3), 603–618. Doi: 10.1017/S003329171200133X

Pinheiro, A. P., Rezaii, M., Rauber, A., Liu, T., Nestor, P. G., McCarley, R. W., … Niznikiewicz, M. A. (2014). Abnormalities in the processing of emotional prosody from single words in schizophrenia. *Schizophrenia Research*, *152*(1), 235–241. Doi: 10.1016/j.schres.2013.10.042

Pinheiro, A. P., Rezaii, N., Nestor, P. G., 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. Doi: 10.1016/j.bandl.2015.12.003

Pinheiro, A. P., Rezaii, N., Rauber, A., & Niznikiewicz, M. (2016). Is this my voice or yours? The role of emotion and acoustic quality in self-other voice discrimination in schizophrenia. *Cognitive Neuropsychiatry*, *21*(4), 1–19. Doi: 10.1080/13546805.2016.1208611

Rosa, C., Lassonde, M., Pinard, C., Keenan, J. P., & Belin, P. (2008). Investigations of hemispheric specialization of self-voice recognition. *Brain and Cognition*, *68*(2), 204–214. Doi: 10.1016/j.bandc.2008.04.007

Rosburg, T., Boutros, N. N., & Ford, J. M. (2008). Reduced auditory evoked potential component N100 in schizophrenia—A critical review. *Psychiatry Research*, *161*(3), 259–274. Doi: 10.1016/j.psychres.2008.03.017

Sambeth, A., Maes, J. H. R., Quiroga, R. Q., & Coenen, A. M. L. (2004). Effects of stimulus repetitions on the event-related potential of humans and rats. *International Journal of Psychophysiology*, *53*(3), 197–205. Doi: 10.1016/j.ijpsycho.2004.04.004

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. Doi: 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. Doi: 10.1016/j.schres.2006.09.021

Stephane, M., Barton, S., & Boutros, N. N. (2001). Auditory verbal hallucinations and dysfunction of the neural substrates of speech. *Schizophrenia Research*, *50*(1–2), 61–78. Doi: 10.1016/S0920-9964(00)00150-X

Stephane, M., Kuskowski, M., McClannahan, K., Surerus, C., & Nelson, K. (2010). Evaluation of speech misattribution bias in schizophrenia. *Psychological Medicine*, *40*(5), 741–748. Doi: 10.1017/S003329170999081X

Swiney, L., & Sousa, P. (2014). A new comparator account of auditory verbal hallucinations: How motor prediction can plausibly contribute to the sense of agency for inner speech. *Frontiers in Human Neuroscience*, *8*, 675. Doi: 10.3389/fnhum.2014.00675

Tacikowski, P., & Nowicka, A. (2010). Allocation of attention to self-name and self-face: An ERP study. *Biological Psychology*, *84*(2), 318–324. Doi: 10.1016/j.biopsycho.2010.03.009

Timm, J., Schönwiesner, M., Schröger, E., & SanMiguel, I. (2016). Sensory suppression of brain responses to self-generated sounds is observed with and without the perception of agency. *Cortex*, *80*, 5–20. Doi: 10.1016/j.cortex.2016.03.018

van der Weiden, A., Prikken, M., & van Haren, N. E. M. (2015). Self-other integration and distinction in schizophrenia: A theoretical analysis and a review of the evidence. *Neuroscience and Biobehavioral Reviews*, *57*, 220–237. Doi: 10.1016/j.neubiorev.2015.09.004

van Lutterveld, R., Sommer, I. E. C., & Ford, J. M. (2011). The neurophysiology of auditory hallucinations—A historical and contemporary review. *Frontiers in Psychiatry*, *2*, 28. Doi: 10.3389/fpsyt.2011.00028

Wang, J., Mathalon, D. H., Roach, B. J., Reilly, J., Keedy, S. K., Sweeney, J. A., & Ford, J. M. (2014). Action planning and predictive coding when speaking. *NeuroImage*, *91*, 91–98. Doi: 10.1016/j.neuroimage.2014.01.003

Waters, F., Allen, P., Aleman, A., Fernyhough, C., Woodward, T. S., Badcock, J. C., … Larøi, F. (2012). Auditory hallucinations in schizophrenia and nonschizophrenia populations: A review and integrated model of cognitive mechanisms. *Schizophrenia Bulletin*, *38*(4), 683–693. Doi: 10.1093/schbul/sbs045

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

Woodruff, P. W. R. (2004). Auditory hallucinations: Insights and questions from neuroimaging. *Cognitive Neuropsychiatry*, *9*(1–2), 73–91. Doi: 10.1080/13546800344000165

## SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

**Figure S1**

**How to cite this article:** Pinheiro AP, Rezaii N, Rauber A, Nestor PG, Spencer KM, Niznikiewicz M. Emotional self–other voice processing in schizophrenia and its relationship with hallucinations: ERP evidence. *Psychophysiology*. 2017;00:1–14. https://doi.org/10.1111/psyp.12880