Schizophrenia Research 152 (2014) 235–241

Contents lists available at ScienceDirect

Schizophrenia Research

journal homepage: www.elsevier.com/locate/schres

Abnormalities in the processing of emotional prosody from single words in schizophrenia

Ana P. Pinheiro [a,b], Neguine Rezaii [b], Andréia Rauber [c], Taosheng Liu [d], Paul G. Nestor [e], Robert W. McCarley [b], Óscar F. Gonçalves [a], Margaret A. Niznikiewicz [b]

[a] Neuropsychophysiology Lab, CIPsi, School of Psychology, University of Minho, Braga, Portugal
[b] Clinical Neuroscience Division, Laboratory of Neuroscience, Department of Psychiatry, Boston VA Healthcare System, Brockton Division and Harvard Medical School, Brockton, MA, United States
[c] Catholic University of Pelotas, Pelotas, Brazil
[d] Department of Psychology, Second Military Medical University (SMMU), Shanghai, China
[e] University of Massachusetts, Boston, MA, United States

ARTICLE INFO

Article history:
Received 4 July 2013
Received in revised form 25 October 2013
Accepted 28 October 2013
Available online 14 December 2013

Keywords:
Schizophrenia
Emotional prosody
Language
Event-related potentials

ABSTRACT

Background: Abnormalities in emotional prosody processing have been consistently reported in schizophrenia and are related to poor social outcomes. However, the role of stimulus complexity in abnormal emotional prosody processing is still unclear.
Method: We recorded event-related potentials in 16 patients with chronic schizophrenia and 16 healthy controls to investigate: 1) the temporal course of emotional prosody processing; and 2) the relative contribution of prosodic and semantic cues in emotional prosody processing. Stimuli were prosodic single words presented in two conditions: with intelligible (semantic content condition—SCC) and unintelligible semantic content (pure prosody condition—PPC).
Results: Relative to healthy controls, schizophrenia patients showed reduced P50 for happy PPC words, and reduced N100 for both neutral and emotional SCC words and for neutral PPC stimuli. Also, increased P200 was observed in schizophrenia for happy prosody in SCC only. Behavioral results revealed higher error rates in schizophrenia for angry prosody in SCC and for happy prosody in PPC.
Conclusions: Together, these data further demonstrate the interactions between abnormal sensory processes and higher-order processes in bringing about emotional prosody processing dysfunction in schizophrenia. They further suggest that impaired emotional prosody processing is dependent on stimulus complexity.

1. Introduction

Among the most significant predictors of long-term disability in schizophrenia (e.g., Couture et al., 2006) is impaired detection and recognition of emotions from voice, i.e., emotional prosody [EP]. Affect recognition from both voice and face is an aspect of social cognition, which has been recently recognized as an important predictor of functional outcomes at all stages of schizophrenia pathology: clinical high risk (Addington et al., 2008; Green et al., 2012), first episode (Horan et al., 2012) and chronic schizophrenia (Kee et al., 2003; Kucharska-Pietura et al., 2005; Green et al., 2012). While face processing abnormality in schizophrenia has been well characterized (e.g., Li et al., 2010), voice and prosody processing have been understudied, especially using event-related potential (ERP) approaches, which remain the only tool to examine temporal changes in neurophysiological events that correspond to early stages of analysis of a speech signal. The existing studies on vocal emotional processing include just a handful of behavioral (e.g., Edwards et al., 2001), functional magnetic resonance imaging (fMRI—e.g., Mitchell et al., 2004; Leitman et al., 2011) and ERP investigations (Pinheiro et al., 2012).

In healthy subjects, perception of emotional prosody is thought to reflect three interacting stages: 1) sensory processing of a speech signal; 2) implicit categorization of salient acoustic features into emotional and non-emotional features; and 3) explicit evaluation and assignment of emotional meaning to a speech signal (Schirmer and Kotz, 2006; Paulmann and Kotz, 2008; Paulmann et al., 2010). Event-related potential (ERP) studies demonstrated that the first two stages are indexed by N100 and P200, respectively (Paulmann and Kotz, 2008; Paulmann et al., 2010; Pinheiro et al., 2012).

Despite the importance of a detailed understanding of emotional prosody processing deficits in schizophrenia, few studies have examined these abnormalities and their underlying neural mechanisms are not well understood. Recent studies suggested that sensory-based dysfunction might not exclusively account for abnormal prosody processing in schizophrenia. Instead, an interaction between dysfunctional sensory and higher-order cognitive processes may better explain it (Leitman et al., 2010, 2011; Pinheiro et al., 2012). A recent ERP study provided further evidence for these abnormalities (Pinheiro et al., 2012). This study investigated prosody processing in 15 chronic schizophrenia patients and 15 healthy controls (HC). Additionally, it explored

0920-9964/$ – see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.schres.2013.10.042

the relative contributions of prosodic and semantic cues. Stimuli were prosodic sentences with intelligible (semantic content condition—SCC) and unintelligible semantic content (pure prosody condition—PPC). The ERP effects occurred within the first 200 ms from the sentence onset in both groups (Pinheiro et al., 2012), supporting previous studies' results (Paulmann and Kotz, 2008; Paulmann et al., 2010). The results revealed abnormalities in the three stages of prosody processing in schizophrenia, which were more pronounced for prosodic SCC sentences. Less negative N100 suggested abnormal sensory processing of prosodic SCC sentences irrespective of valence. Increased P200 to angry and happy prosodic stimuli in the SCC, and to happy stimuli in the PPC suggested abnormal detection of emotional salience. Behavioral results revealed impaired cognitive evaluation of the emotional significance of angry SCC and neutral PPC sentences.

In view of a critical need for a systematic study of emotional prosody processing in schizophrenia, the current study extended our previous work, by investigating the temporal course of prosody processing using *single words* with both intelligible (SCC) and unintelligible semantic content (PPC). Based on language studies demonstrating differences in the processing of words in a sentence vs. in isolation (e.g., Van Petten, 1995) and effects of phrasal length and complexity on prosodic processing (Wheeldon and Lahiri, 1997; Krivokapi, 2007), we reasoned that prosody processing of sentences may differ from that of single words. For example, the processing of words embedded in a sentence is susceptible to syntactic and semantic constraints imposed by a sentence context, which can modify many aspects of their processing (e.g., Van Petten, 1995). Furthermore, in relation to words in isolation, the processing of a sentence demands more working memory and attention resources, as meaning is built up across the course of the sentence (e.g., Van Petten, 1995). Thus, considering the attentional (Nestor et al., 2001; Laurens et al., 2005) and verbal working memory deficits (Menon et al., 2001; Silver et al., 2003) often reported in schizophrenia, the processing of prosodic information may be more impaired in sentences than in single words.

Because of its excellent temporal resolution, we used ERPs to address the role of stimulus complexity in the first two stages of emotional prosody processing: the sensory processing of prosodic information (N100) and the detection of its emotional salience (P200), both processes not accessible to behavioral probes. We also collected data on accuracy of prosody recognition to shed light on a later stage of emotional prosody processing, i.e. the assignment of emotional meaning to a voice signal. We hypothesized that if impaired prosody processing is not dependent on stimulus complexity, similar abnormalities to those reported in Pinheiro et al. (2012) will be observed in the current study. However, if stimulus complexity matters, we expected less severe prosody processing abnormalities in the single word relative to the sentence prosody processing study.

Considering previous studies demonstrating an association between deficits in emotional prosody recognition and positive symptomatology (Poole et al., 2000; Rossell and Boundy, 2005; Shea et al., 2007), and between increased P200 amplitude for happy prosody and delusions (Pinheiro et al., 2012), we predicted that ERP abnormalities amplitude would be associated with positive symptomatology scores.

2. Method

2.1. Participants

Sixteen patients with a diagnosis of chronic schizophrenia and 17 HC matched for age, handedness and parental socioeconomic status (Hollingshead, 1976) participated in this study (Table 1). Subjects had normal hearing as assessed by audiometry, and normal or corrected to normal vision. Patients were recruited at the Veterans Affairs Hospital, Brockton and HC were recruited from Internet advertisements.

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) in the last year prior to EEG assessment; full scale intelligence quotient (IQ) above 85 (Wechsler, 2008); no hearing, vision or upper body impairment. For HC, additional inclusion criteria were: no history of Axis I–II disorders (First et al., 1995, 2002); no history of Axis I disorder in first or second-degree relatives (Andreasen et al., 1977).

Patients were diagnosed (screened for HC) using the SCID-I and SCID-II (First et al., 1995, 2002). Symptom severity was assessed with the *Positive and Negative Syndrome Scale* (PANSS-Kay et al., 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.

Table 1
Demographic and clinical characteristics of participants.

Variable	Healthy controls (n = 17)	Schizophrenia patients (n = 16)	*p* value[a]
Age (years)	48.13 ± 5.66	48.86 ± 7.40	.750
Women, *n*	7	5	
Education (years)	15.18 ± 1.64	14.00 ± 2.42	.119
Subject's SES[b]	2.13 ± 0.81	2.93 ± 1.14	.033[*]
Parental SES	2.44 ± 0.81	2.79 ± 1.53	.434
Handedness[c]	0.81 ± 0.15	0.79 ± 0.21	.848
Neurocognitive data			
Full scale composite score	99.33 ± 12.30	92.79 ± 14.32	.227
Verbal comprehension composite score	99.08 ± 11.47	95.93 ± 15.82	.572
Working memory composite score	105.33 ± 14.22	92.86 ± 12.90	.049[*]
Processing speed composite score	101.17 ± 89.64	89.64 ± 14.87	.107
Clinical data			
Onset age (years)	NA	30.07 ± 11.23	NA
Duration (years)	NA	19.47 ± 10.95	NA
Chlorpromazine EQ (mg)	NA	356.78 ± 294.56	NA
Antipsychotic medication type	NA	Typical (fluphenazine decanoate, proloxin decanoate, haloperidol) = 3; Atypical (risperidone, olanzapine, ziprasidone, quetiapine, aripiprazole) = 11	NA
Other psychotropic medication	NA	Antidepressants (sertraline, citalopram, bupoprion, trazodone) = 4 Benzodiazepines (lorazepam, clonazepam) = 4 Lithium carbonate = 2 Valproic acid = 3	NA
PANSS delusions	NA	4.88 ± 2.16	NA
PANSS conceptual disorganization	NA	2.50 ± 1.10	NA
PANSS hallucinations	NA	4.00 ± 2.19	NA
PANSS positive scale	NA	20.25 ± 8.19	NA
PANSS negative scale	NA	22.88 ± 9.76	NA
PANSS general psychopathology	NA	38.56 ± 11.70	NA
PANSS total psychopathology	NA	81.69 ± 25.92	NA
SANS total	NA	10.59 ± 5.44	NA
SAPS total	NA	9.63 ± 3.05	NA

Notes. All values represent mean ± SD. SES = socioeconomic status; Chlorpromazine EQ = chlorpromazine equivalent dose; NA = not applicable.

[a] Independent samples *t*-test tested for group differences in sociodemographic and neurocognitive measures.

[b] Hollingshead Four-Factor Index of Social Status (Hollingshead, 1976).

[c] Edinburgh Handedness Inventory (Oldfield, 1971).

[*] $p < 0.05$.

2.2. Stimuli

Stimuli used in the SCC were 40 words with neutral semantic content and short length (e.g., “card”, “pen”—see Supplementary Material). Words were controlled for frequency ($M = 10.38 \pm 11.05$), familiarity ($M = 582.37 \pm 25.88$), age of acquisition ($M = 232.54 \pm 55.15$), concreteness ($M = 594.60 \pm 43.79$), and number of letters ($M = 5.14 \pm 1.81$), based on the MRC Psycholinguistic Database (Coltheart, 1981; Wilson, 1998).

Words were recorded by an American English native speaker with training in theater techniques, with neutral and emotional prosody (happy; angry) using an Edirol R-09 recorder and a CS-15 cardioid-type stereo microphone, at a sampling rate of 22 kHz and 16-bit quantization. Words' pitch, intensity and duration were compared across conditions (Table 2). Duration of happy words was longer than duration of angry ($p < 0.01$) and neutral words ($p < 0.001$). Mean pitch was higher for happy relative to both angry ($p < 0.001$) and neutral words ($p < 0.001$), and for angry relative to neutral words ($p < 0.001$). Mean intensity did not differ across emotion types ($p > 0.05$).

Fifteen subjects (7 female) who did not participate in the ERP sessions assessed the valence of words' intonation. Angry words were rated as “angry” by 96.11%, happy words were rated as “happy” by 99.18%, and neutral words were rated as “neutral” by 96.66% of participants.

In PPC, the same stimuli were distorted to make their semantic content unintelligible by using Praat software (Boersma and Weenink, 2006). The phones of each SCC word were manually replaced by phones produced by the same speaker, preserving both the original voice and prosodic features. All fricatives were replaced with the phone [s], all stop consonants with [t], all glides with [j], all stressed vowels with [æ], and all unstressed vowels with [ə].

2.3. Procedure

Each participant was seated comfortably at a distance of 100 cm from a computer monitor in a sound-attenuating chamber. The experimental session was divided into two blocks (block1: SCC words; block2: PPC words). Block order was counterbalanced. Each block contained 105 words of different prosody types (35 neutral, 35 happy, 35 angry). The remaining five words of each valence and type (SCC, PPC), from the original list of 40 words, were presented in the practice block. Stimuli were presented binaurally through headphones at a comfortable sound level. Superlab Pro software package (2008) controlled stimulus delivery.

Before each experimental block, participants were given a brief training with feedback. Fig. 1 illustrates an experimental trial. Before each word onset, a fixation cross was presented centrally on the screen for 1000 ms, and was kept during word presentation to minimize eye movements. After 1000 ms, a question mark signaled the beginning of the response time (5 s). Subjects were instructed to make a decision whether a word was spoken with a neutral, happy, or angry tone of voice by pressing one of the three buttons. The order of button presses was counterbalanced across subjects. Each response key was marked with an emoticon to minimize working memory demands. A 2000 ms inter-stimulus interval separated the end of an event and the beginning of the next one. A short pause was provided after 15 words. During the experiment, no feedback was provided.

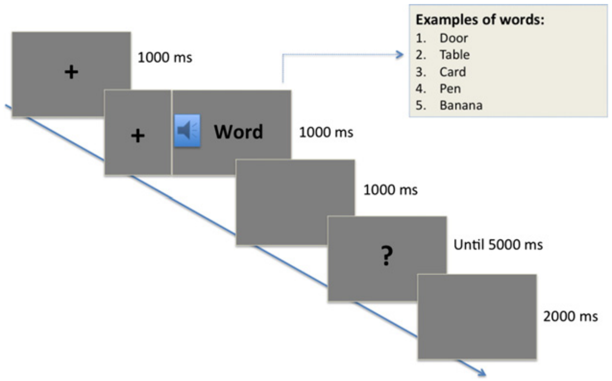


Fig. 1. Illustration of an experimental trial.

2.4. Data acquisition and analysis

EEG was recorded with custom-made electrode caps with a 64-channel BioSemi Active-Two system (BioSemi B.V., The Netherlands). It was acquired in a continuous mode at a digitization rate of 512 Hz, with a bandpass of 0.01–100 Hz. Blinks and eye movements were monitored through electrodes placed on the left and right temples and below the left eye.

EEG data were processed offline using Brain Analyzer 2 package (Brain Products, Germany), and re-referenced offline to the mathematical average of the left and right mastoids. Individual ERP epochs were created for each prosody type (neutral, happy, angry) in each word condition (SCC, PPC), with −200 ms pre-stimulus baseline and 900 ms post-stimulus epoch. Eye blinks and movement artifacts were corrected by the Gratton et al. (1983) method. EEG epochs containing muscle activity or amplifier blocking were rejected offline before averaging (+/−100 μV criterion). After artifact rejection, at least 75% of trials per condition per subject entered the analyses. The number of individual trials did not differ between groups ($p > .05$).

The inspection of grand average waveforms (Figs. 2 and 3) revealed three main components with predominantly central distribution: a positivity occurring around 50 ms (P50), a negativity occurring around 100 ms (N100), and a positivity occurring around 200 ms (P200). Temporal windows were then selected for P50, N100 and P200 based on the visual inspection of the waveforms. Mean amplitude was calculated between 30 and 125 ms (P50), 125 and 190 ms (N100), and 220 and 320 ms (P200), post-stimulus onset, at central electrodes (Cz, C3, C4).

Table 2
Acoustic properties of words with angry, happy and neutral prosody in the semantic content (SCC) and pure prosody (PPC) conditions.

Semantic status	Emotion	Acoustic properties						
		Duration	F0			Intensity		
			Min	Mean	Max	Min	Mean	Max
SCC	Angry	639.00 (19.45)	163.09 (2.13)	271.94 (4.12)	368.15 (5.68)	45.67 (1.05)	78.64 (0.39)	86.24 (0.28)
	Happy	718.21 (19.89)	177.58 (11.44)	371.03 (12.45)	574.94 (13.53)	50.58 (0.84)	79.27 (0.30)	86.49 (0.26)
	Neutral	663.23 (22.12)	156.91 (2.04)	183.70 (1.24)	228.09 (4.36)	49.82 (1.41)	78.97 (0.28)	85.09 (0.22)
PPC	Angry	653.94 (25.65)	170.55 (5.15)	265.00 (4.38)	363.09 (6.98)	43.18 (0.83)	77.85 (0.36)	84.88 (0.39)
	Happy	722.27 (24.34)	173.03 (8.76)	377.73 (10.01)	559.00 (12.91)	48.27 (1.05)	78.73 (0.28)	86.00 (0.26)
	Neutral	629.24 (28.83)	156.70 (1.59)	181.55 (1.52)	219.09 (2.55)	48.39 (1.68)	77.82 (0.35)	83.79 (0.22)

Notes. Mean (standard error); min = minimum; max = maximum.

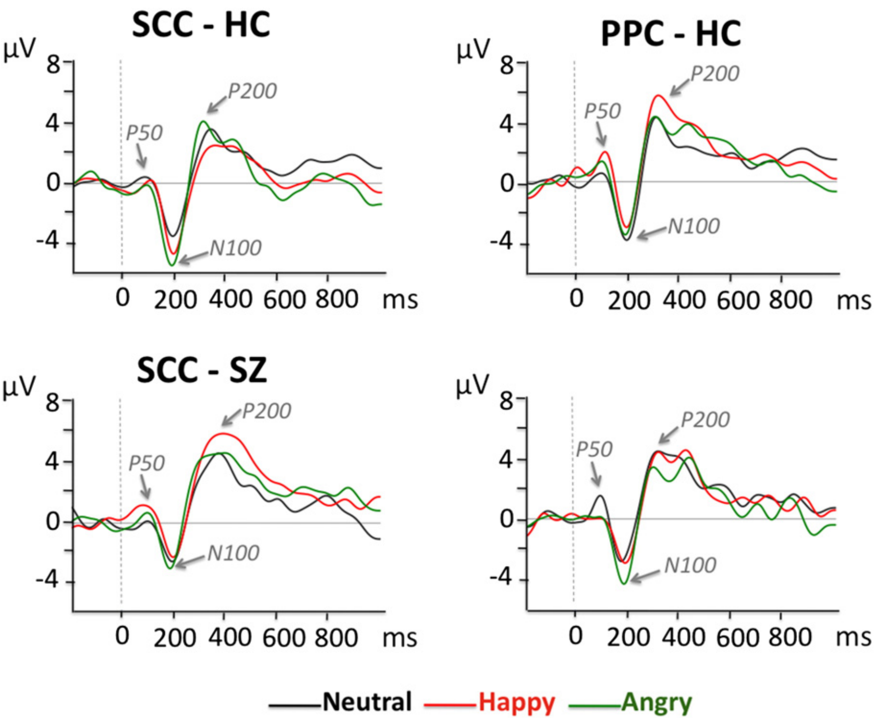


Fig. 2. Grand average waveforms for neutral, happy and angry prosody in the semantic content condition (SCC) and pure prosody condition (PPC) at Cz, in healthy controls (HC) and schizophrenia patients (SZ).

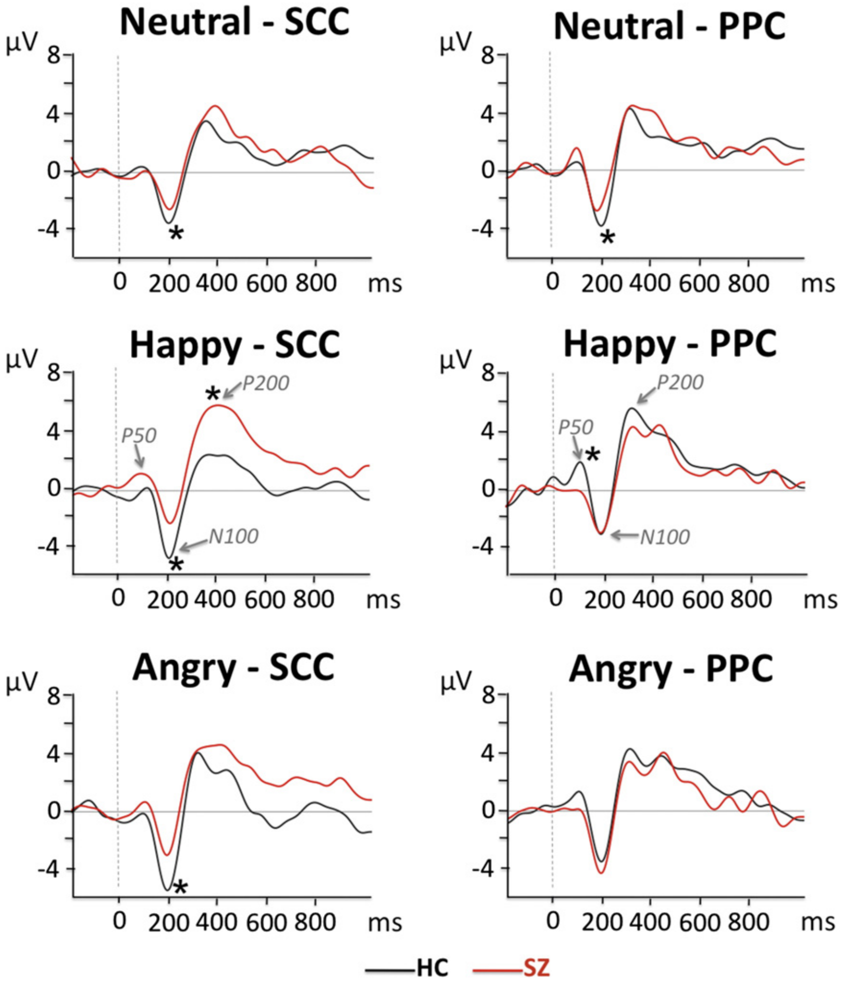


Fig. 3. Grand average waveforms showing group contrasts for neutral, happy, and angry prosody in the semantic content condition (SCC) and pure prosody condition (PPC) at Cz.

2.5. Statistical analyses

For the statistical analysis, the PAWS 20.00 (SPSS Inc., USA) software package was used. Only significant results are presented ($p < 0.05$).

2.5.1. ERP data

Repeated measures analyses of variance (ANOVAs) were computed for the between-group comparisons of N100 and P200 peak amplitude, with semantic status (SCC, PPC), emotion (neutral, happy, angry), and electrodes (Cz, C3, C4) as within-subject factors and group as a between-subject factor, using IBM SPSS Statistics 20 (SPSS Inc., USA).

2.5.2. Accuracy data

A repeated measures ANOVA with semantic status and emotion as within-subjects factors and group as between-subjects factor tested group differences in behavioral accuracy.

Analyses were corrected for non-sphericity using the Greenhouse–Geisser method (the original *df* is reported). All significance levels are two-tailed with the preset significance alpha level of $p < 0.05$. Main effects were followed with pairwise comparisons between conditions, using the Bonferroni adjustment for multiple comparisons.

3. Results

3.1. ERP data (Figs. 2 and 3)

3.1.1. P50 amplitude

A significant group × semantic status × emotion interaction was observed ($F(2, 62) = 6.603$, $p < 0.01$). We followed-up this interaction with subsequent ANOVAs for each semantic status condition separately. A significant group × emotion interaction was observed in the PPC only ($F(2, 62) = 4.872$, $p = 0.016$): groups differed in the processing of happy PPC prosody ($p < 0.01$), with reduced P50 amplitude in patients relative to HC.

3.1.2. N100 amplitude

A main effect of emotion ($F(2, 62) = 6.723$, $p < 0.01$) revealed that N100 was more negative for angry relative to neutral prosody ($p < 0.01$) and tended to be more negative for angry relative to happy prosody ($p = 0.083$) in both groups. A significant group × semantic status × emotion interaction ($F(2, 62) = 4.638$, $p = 0.02$) indicated differences in the way patients and HC processed prosodic stimuli at the sensory level.

Follow-up separate repeated-measures ANOVAs for each semantic status condition showed a significant group effect for the SCC ($F(1, 31) = 8.395$, $p < 0.01$): N100 was overall less negative in the schizophrenia group relative to HC. In addition, a significant group × emotion interaction was observed in the PPC ($F(2, 62) = 3.874$, $p = 0.027$). Pairwise comparisons indicated less negative N100 in schizophrenia relative to HC subjects in the neutral condition only ($p = 0.039$).

3.1.3. P200 amplitude

A significant group × semantic status × emotion interaction ($F(2, 62) = 5.476$, $p < 0.01$) indicated differences in the way groups integrated acoustic information into an emotional percept. Separate ANOVAs were subsequently computed for each semantic status condition. A significant group × emotion interaction was observed for SCC ($F(2, 62) = 3.215$, $p = 0.049$). Subsequent pairwise comparisons indicated more positive P200 for happy prosody in patients relative to HC ($p = 0.01$). No significant effects were observed for PPC.

3.2. Accuracy data

More correct responses were found in SCC relative to PPC (main effect of semantic status—$F(1, 31) = 45.606$, $p < 0.001$). A significant group × semantic status × emotion interaction ($F(2, 62) = 4.327$, $p = 0.020$) indicated more incorrect responses for angry SCC words ($p = 0.036$) and happy PPC words ($p = 0.029$) in schizophrenia (Fig. 4). However, no main effect of group was observed ($p > 0.05$).

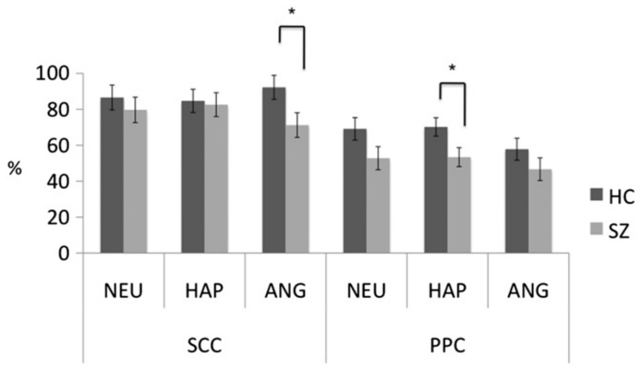
3.3. Correlational analyses

Two-tailed Spearman's rho correlation analyses were conducted in an exploratory analysis of the relationship between schizophrenia abnormalities in P50 (happy PPC), N100 (neutral, happy, and angry SCC; neutral PPC) and P200 (happy SCC) amplitude at Cz and: 1) clinical symptoms (PANSS), medication (chlorpromazine equivalent) and illness duration; 2) neurocognitive data (WAIS composite scores); and 3) behavioral indices of prosody recognition. The significance level was adjusted for multiple comparisons using Bonferroni correction. No significant correlations were found ($p > 0.05$).

4. Discussion

This study extended and clarified our previous findings for prosodic sentences (Pinheiro et al., 2012). ERP and behavioral findings showed group differences that spanned the three stages of prosody processing and interacted with the semantic status of words. ERP effects were observed within the first 200 ms. In addition to N100 and P200, we observed prosodic effects in an earlier time window around 50 ms (P50), corroborating the sensitivity of P50, N100 and P200 components to prosodic manipulations in speech sounds (Paulmann and Kotz, 2008; Paulmann et al., 2010; Liu et al., 2012; Pinheiro et al., 2012). Schizophrenia patients showed a markedly different P50, N100 and P200 pattern as a function of both semantic status and emotion type relative to HC.

Reduced P50 amplitude for happy PPC prosody was observed in schizophrenia relative to HC. P50 has been reported in studies of auditory gating (e.g., Boutros et al., 2004) and has been considered an index of the formation of sensory memory traces at the level of the primary auditory cortex (Haenschel et al., 2005). The existing evidence suggests that P50 amplitude may be modulated by the physical properties of the eliciting stimulus (Chen et al., 1997; Ninomiya et al., 2000) and by attention (Erwin et al., 1998). Also, in our previous study (Liu et al., 2012) with non-verbal vocalizations, emotion effects were found at the level of P50 (Liu et al., 2012). Reduced P50 amplitude has been consistently demonstrated in schizophrenia (e.g., Potter et al., 2006). In our study, reduced P50 for happy PPC in schizophrenia points to abnormal early somatosensory information processing that is stimulus specific.



	SCC			PPC		
	Neutral	Happy	Angry	Neutral	Happy	Angry
HC	86.57 (6.89)	84.68 (6.46)	92.21 (6.68)	69.11 (6.25)	70.18 (5.11)	57.79 (6.11)
SZ	79.69 (7.11)	82.59 (6.64)	71.21 (6.86)	52.80 (6.43)	53.39 (5.29)	46.65 (6.32)

Fig. 4. Percentage of correct responses in the recognition of emotional prosody in both semantic content condition (SCC) and pure prosody condition (PPC) in healthy controls (HC) and schizophrenia patients (SZ).

Reduced N100 amplitude in schizophrenia was found to both emotional and neutral SCC stimuli, as well as to neutral PPC stimuli. The N100 component is related to early auditory encoding, and its amplitude is modulated by the physical properties of the stimuli and by allocation of attentional resources (Rosburg et al., 2008). P50 and N1 are thought to represent distinct aspects of information processing (Boutros et al., 2004). Considering the functional role of N100 as an index of initial sensory processing of the prosodic signal (Schirmer and Kotz, 2006), these findings support deficits in sensory processing of vocal information (e.g., Pinheiro et al., 2012) that were enhanced when semantic information was present. Given that N100 generators are located mainly in supratemporal plane and superior temporal gyrus (Naatanen and Picton, 1987), reduced N100 amplitude may reflect functional and structural brain changes in temporal structures that are a central feature of the schizophrenia diagnosis (e.g., Shenton et al., 1992). However, since N100 cannot be directly related to a single cortical process and is influenced by many individual-related variables, we cannot rule out the contributions of other factors, such as attention or arousal.

Specific abnormalities were noted in the second stage of prosody processing as indexed by increased P200 amplitude to happy SCC words only in schizophrenia relative to HC. The P200 is primarily generated in the temporal cortex (such as the planum temporale and the auditory association complex, area 22—Godey et al., 2001), even though frontal areas are also involved (McCarley et al., 1991). Considering the role of P200 as an index of the emotional salience of a vocal stimulus (Paulmann and Kotz, 2008), as proposed in the multi-stage model of emotional prosody processing (Schirmer and Kotz, 2006), increased P200 for happy prosody might indicate a specific impairment in categorizing happy auditory emotional percepts as “salient”. However, this was the case only when happy prosodic information was embedded in intelligible speech suggesting that sensory cues were used differently in the two conditions. Additionally, given the sensitivity of P200 to task difficulty (increased P200 amplitude related to increased cognitive effort—Lenz et al., 2007), it is plausible that the salience of positive social information was more difficult to extract for schizophrenia patients. Also, given that all stimuli had neutral semantic content but could carry emotional intonation, we cannot rule out the effects of incongruity (semantic vs. prosodic) on P200 amplitude (Scholten et al., 2008).

Two major conclusions arise from P50, N100 and P200 findings in the current study: 1) the fact that group differences were not observed for all types of prosodic stimuli speaks against a generalized prosodic impairment and suggests that prosodic abnormalities may be dependent on stimulus type; 2) abnormalities in the processing of emotional but not neutral cues seem to be more pronounced when speech's semantic content is intelligible, suggesting that abnormalities in the processing of both semantic and prosodic aspects of voice interact since early stages of prosody processing.

Behavioral data, indexing the integration of emotionally significant acoustic cues (Schirmer and Kotz, 2006), indicated that emotional prosody recognition was better in SCC relative to PPC in both groups, confirming our initial hypothesis (and also Pinheiro et al., 2012). Given the absence of a memory representation for unintelligible stimuli to facilitate predictive processes, this result likely reflects increased task demands. Additionally, schizophrenia patients made more errors in identifying *emotional* but not neutral prosody. This result suggests that deficits in recognizing emotional prosody in single words depend both on emotion type and on semantic status.

Finally, we note differences in both ERP and behavioral results reported in this and in our previous study using prosodic sentences (Pinheiro et al., 2012). In our previous study, P200 was increased to happy SCC and PPC sentences and to angry SCC sentences; here, P200 abnormalities were observed only for happy SCC words. These ERP differences suggest that while processing prosody in sentences and words evokes similar ERP components, the processes involved are not identical. They additionally suggest that stimulus complexity may differently impact sensory and early categorization stages of prosody processing.

Furthermore, the overall reduced emotional recognition accuracy observed in the sentences study contrasted with specific deficits in the recognition of angry SCC prosody and happy PPC prosody in single words. Since prosody processing relies on the continuous monitoring of dynamically changing acoustic cues underlying an emotional tone, greater working memory and attention demands exist for sentences vs. single words. Accordingly, they were associated with more errors in identifying sentence relative to single word prosodic stimuli.

4.1. Limitations

Limitations of this study are a sample composed by medicated chronic schizophrenia patients. Future research with unmedicated and first-episode patients will overcome some of the limitations associated with medication and chronicity.

Role of funding source

This work was supported by two grants awarded to A.P.P.: Post-Doctoral Grant no. SFRH/BD/35882/2007 funded by *Fundação para a Ciência e a Tecnologia* (FCT, Portugal); and research grant no. PTDC/PSI-PCL/116626/2010, funded by FCT and FEDER (*Fundo Europeu de Desenvolvimento Regional*) through the European programs QREN (*Quadro de Referência Estratégico Nacional*) and COMPETE (*Programa Operacional Factores de Competitividade*). It was additionally supported by two grants from the National Institute of Mental Health (no. RO1 MH 040799 awarded to R.W.M. and no. RO3 MH 078036 awarded to M.A.N.).

Contributors

Ana P. Pinheiro and Margaret Niznikiewicz designed the study and wrote the protocol. Ana P. Pinheiro, Margaret Niznikiewicz, Neguine Rezaii and Taosheng Liu collected and analyzed the data. All authors collaborated in the statistical analysis. Andréia Rauber edited the stimuli and completed the acoustic analyses of the stimuli. Paul G. Nestor did the clinical and neuropsychological testing of participants. Ana P. Pinheiro and Margaret Niznikiewicz wrote the first draft of the manuscript. All authors contributed to and have approved the final manuscript.

Conflict of interest

All authors report no competing interests.

Acknowledgments

We are grateful to all the participants of this study for their contribution to science. We are also grateful to Elizabeth Thompson and Israel Molina for their help with data acquisition, and to Victoria Choate for her invaluable help with voice recording.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.schres.2013.10.042.

References

Addington, J., Penn, D., Woods, S.W., Addington, D., Perkins, D.O., 2008. Facial affect recognition in individuals at clinical high risk for psychosis. Br. J. Psychiatry 192 (1), 67–68.

American Psychiatric Association, 2000. Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR. American Psychiatric Association Press, Washington, DC.

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

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

Andreasen, N.C., Endicott, J., Spitzer, R.L., Winokur, G., 1977. The family history method using diagnostic criteria: reliability and validity. Arch. Gen. Psychiatry 34, 1229–1235.

Boersma, P., Weenink, D., 2006. Praat: Doing Phonetics by Computer (Version 5.0.43) [Computer Program].

Boutros, N.N., Korzyukov, O., Jansen, B., Feingold, A., Bell, M., 2004. Sensory gating deficits during the mid-latency phase of information processing in medicated schizophrenia patients. Psychiatry Res. 126 (3), 203–215.

Chen, C.H., Ninomiya, H., Onitsuka, T., 1997. Influence of reference electrodes, stimulation characteristics and task paradigms on auditory P50. Psychiatry Clin. Neurosci. 51 (3), 139–143.

Coltheart, M., 1981. The MRC psycholinguistics database. Q. J. Exp. Psychol. A 33A, 497–505.

Couture, S.M., Penn, D.L., Roberts, D.L., 2006. The functional significance of social cognition in schizophrenia: a review. Schizophr. Bull. 32 (Suppl. 1), S44–S63.

Edwards, J., Pattison, P.E., Jackson, H.J., Wales, R.J., 2001. Facial affect and affective prosody recognition in first-episode schizophrenia. Schizophr. Res. 48 (2–3), 235–253.

Erwin, R.J., Turetsky, B.I., Moberg, P., Gur, R.C., Gur, R.E., 1998. P50 abnormalities in schizophrenia: relationship to clinical and neuropsychological indices of attention. Schizophr. Res. 33 (3), 157–167.

First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.M., 1995. Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II, Version 2.0). Biometrics Research Department, New York State Psychiatric Institute, New York, NY.

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). Biometric Research Department, New York State Psychiatric Institute, New York, NY.

Godey, B., Schwartz, D., de Graaf, J.B., Chauvel, P., Liegeois-Chauvel, C., 2001. Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: a comparison of data in the same patients. Clin. Neurophysiol. 112 (10), 1850–1859.

Gratton, G., Coles, M.G., Donchin, E., 1983. A new method for off-line removal of ocular artifact. Electroencephalogr. Clin. Neurophysiol. 55 (4), 468–484.

Green, M.F., Bearden, C.E., Cannon, T.D., Fiske, A.P., Hellemann, G.S., Horan, W.P., Kee, K., Kern, R.S., Lee, J., Sergi, M.J., Subotnik, K.L., Sugar, C.A., Ventura, J., Yee, C.M., Nuechterlein, K.H., 2012. Social cognition in schizophrenia, part 1: performance across phase of illness. Schizophr. Bull. 38 (4), 854–864.

Haenschel, C., Vernon, D.J., Dwivedi, P., Gruzelier, J.H., Baldeweg, T., 2005. Event-related brain potential correlates of human auditory sensory memory-trace formation. J. Neurosci. 25 (45), 10494–10501.

Hollingshead, A., 1976. Four Factor Index of Social Status. Yale University Social Sciences Library, New Haven, CT.

Horan, W.P., Green, M.F., DeGroot, M., Fiske, A., Hellemann, G., Kee, K., Kern, R.S., Lee, J., Sergi, M.J., Subotnik, K.L., Sugar, C.A., Ventura, J., Nuechterlein, K.H., 2012. Social cognition in schizophrenia, part 2: 12-month stability and prediction of functional outcome in first-episode patients. Schizophr. Bull. 38 (4), 865–872.

Kay, S.R., Fiszbein, A., Opler, L.A., 1987. The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. Schizophr. Bull. 13 (2), 261–276.

Kee, K.S., Green, M.F., Mintz, J., Brekke, J.S., 2003. Is emotion processing a predictor of functional outcome in schizophrenia? Schizophr. Bull. 29 (3), 487–497.

Krivokapi, J., 2007. Prosodic planning: effects of phrasal length and complexity on pause duration. J. Phon. 35 (2), 162–179.

Kucharska-Pietura, K., David, A., Masiak, K., Phillips, M., 2005. Perception of facial and vocal affect by people with schizophrenia in early and late stages of illness. Br. J. Psychiatry 187, 523–528.

Laurens, K.R., Kiehl, K.A., Ngan, E.T., Liddle, P.F., 2005. Attention orienting dysfunction during salient novel stimulus processing in schizophrenia. Schizophr. Res. 75 (2–3), 159–171.

Leitman, D.I., Laukka, P., Juslin, P.N., Saccente, E., Butler, P., Javitt, D.C., 2010. Getting the cue: sensory contributions to auditory emotion recognition impairments in schizophrenia. Schizophr. Bull. 36 (3), 545–556.

Leitman, D.I., Wolf, D.H., Laukka, P., Ragland, J.D., Valdez, J.N., Turetsky, B.I., Gur, R.E., Gur, R.C., 2011. Not pitch perfect: sensory contributions to affective communication impairment in schizophrenia. Biol. Psychiatry 70 (7), 611–618.

Lenz, D., Schadow, J., Thaerig, S., Busch, N.A., Herrmann, C.S., 2007. What's that sound? Matches with auditory long-term memory induce gamma activity in human EEG. Int. J. Psychophysiol. 64 (1), 31–38.

Li, H., Chan, R.C., McAlonan, G.M., Gong, Q.Y., 2010. Facial emotion processing in schizophrenia: a meta-analysis of functional neuroimaging data. Schizophr. Bull. 36 (5), 1029–1039.

Liu, T., Pinheiro, A.P., Deng, G., Nestor, P.G., McCarley, R.W., Niznikiewicz, M.A., 2012. Electrophysiological insights into processing nonverbal emotional vocalizations. Neuroreport 23 (2), 108–112.

McCarley, R.W., Faux, S.F., Shenton, M.E., Nestor, P.G., Adams, J., 1991. Event-related potentials in schizophrenia: their biological and clinical correlates and a new model of schizophrenic pathophysiology. Schizophr. Res. 4 (2), 209–231.

Menon, V., Anagnoson, R.T., Mathalon, D.H., Glover, G.H., Pfefferbaum, A., 2001. Functional neuroanatomy of auditory working memory in schizophrenia: relation to positive and negative symptoms. Neuroimage 13 (3), 433–446.

Mitchell, R.L., Elliott, R., Barry, M., Cruttenden, A., Woodruff, P.W., 2004. Neural response to emotional prosody in schizophrenia and in bipolar affective disorder. Br. J. Psychiatry 184, 223–230.

Naatanen, R., Picton, T., 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. Psychophysiology 24 (4), 375–425.

Nestor, P.G., Han, S.D., Niznikiewicz, M., Salisbury, D., Spencer, K., Shenton, M.E., McCarley, R.W., 2001. Semantic disturbance in schizophrenia and its relationship to the cognitive neuroscience of attention. Biol. Psychol. 57 (1–3), 23–46.

Ninomiya, H., Sato, E., Onitsuka, T., Hayashida, T., Tashiro, N., 2000. Auditory P50 obtained with a repetitive stimulus paradigm shows suppression to high-intensity tones. Psychiatry Clin. Neurosci. 54 (4), 493–497.

Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9 (1), 97–113.

Paulmann, S., Kotz, S.A., 2008. Early emotional prosody perception based on different speaker voices. Neuroreport 19 (2), 209–213.

Paulmann, S., Seifert, S., Kotz, S.A., 2010. Orbito-frontal lesions cause impairment during late but not early emotional prosodic processing. Soc. Neurosci. 5 (1), 59–75.

Pinheiro, A.P., Del Re, E., Mezin, J., Nestor, P.G., Rauber, A., McCarley, R.W., Goncalves, O.F., Niznikiewicz, M.A., 2012. Sensory-based and higher-order operations contribute to abnormal emotional prosody processing in schizophrenia: an electrophysiological investigation. Psychol. Med. 1–16.

Poole, J.H., Tobias, F.C., Vinogradov, S., 2000. The functional relevance of affect recognition errors in schizophrenia. J. Int. Neuropsychol. Soc. 6 (6), 649–658.

Potter, D., Summerfelt, A., Gold, J., Buchanan, R.W., 2006. Review of clinical correlates of P50 sensory gating abnormalities in patients with schizophrenia. Schizophr. Bull. 32 (4), 692–700.

Rosburg, T., Boutros, N.N., Ford, J.M., 2008. Reduced auditory evoked potential component N100 in schizophrenia—a critical review. Psychiatry Res. 161 (3), 259–274.

Rossell, S.L., Boundy, C.L., 2005. Are auditory-verbal hallucinations associated with auditory affective processing deficits? Schizophr. Res. 78 (1), 95–106.

Schirmer, A., Kotz, S.A., 2006. Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. Trends Cogn. Sci. 10 (1), 24–30.

Scholten, M.R., Aleman, A., Kahn, R.S., 2008. The processing of emotional prosody and semantics in schizophrenia: relationship to gender and IQ. Psychol. Med. 38 (6), 887–898.

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. Schizophr. Res. 90, 214–220.

Shenton, M.E., Kikinis, R., Jolesz, F.A., Pollak, S.D., LeMay, M., Wible, C.G., Hokama, H., Martin, J., Metcalf, D., Coleman, M., et al., 1992. Abnormalities of the left temporal lobe and thought disorder in schizophrenia. A quantitative magnetic resonance imaging study. N. Engl. J. Med. 327 (9), 604–612.

Silver, H., Feldman, P., Bilker, W., Gur, R.C., 2003. Working memory deficit as a core neuropsychological dysfunction in schizophrenia. Am. J. Psychiatry 160 (10), 1809–1816.

van Petten, C., 1995. Words and sentences: event-related brain potential measures. Psychophysiology 32 (6), 511–525.

Wechsler, D., 2008. Wechsler Adult Intelligence Scale: Administration and Scoring Manual, 4th edition. The Psychological Corporation, San Antonio.

Wheeldon, L., Lahiri, A., 1997. Prosodic units in speech production. J. Mem. Lang. 37 (3), 356–381.

Wilson, M.D., 1998. The MRC psycholinguistic database: machine readable dictionary, version 2. Behav. Res. Methods Instrum. 20, 6–11.