Schizophrenia Research 143 (2013) 348–353

Contents lists available at SciVerse ScienceDirect

Schizophrenia Research

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

Relationship between auditory processing and affective prosody in schizophrenia

Carol Jahshan *, Jonathan K. Wynn, Michael F. Green

Sierra Pacific Mental Illness Research, Education and Clinical Center (MIRECC), VA Greater Los Angeles Healthcare System, Los Angeles, CA, USA
Department of Psychiatry and Behavioral Sciences, David Geffen School of Medicine, University of California, Los Angeles, CA, USA

ARTICLE INFO

Article history:
Received 21 September 2012
Received in revised form 16 November 2012
Accepted 20 November 2012
Available online 29 December 2012

Keywords:
Schizophrenia
Auditory processing
Affective prosody
Mismatch negativity
P300

ABSTRACT

Patients with schizophrenia have well-established deficits in their ability to identify emotion from facial expression and tone of voice. In the visual modality, there is strong evidence that basic processing deficits contribute to impaired facial affect recognition in schizophrenia. However, few studies have examined the auditory modality for mechanisms underlying affective prosody identification. In this study, we explored links between different stages of auditory processing, using event-related potentials (ERPs), and affective prosody detection in schizophrenia. Thirty-six schizophrenia patients and 18 healthy control subjects received tasks of affective prosody, facial emotion identification, and tone matching, as well as two auditory oddball paradigms, one passive for mismatch negativity (MMN) and one active for P300. Patients had significantly reduced MMN and P300 amplitudes, impaired auditory and visual emotion recognition, and poorer tone matching performance, relative to healthy controls. Correlations between ERP and behavioral measures within the patient group revealed significant associations between affective prosody recognition and both MMN and P300 amplitudes. These relationships were modality specific, as MMN and P300 did not correlate with facial emotion recognition. The two ERP waves accounted for 49% of the variance in affective prosody in a regression analysis. Our results support previous suggestions of a relationship between basic auditory processing abnormalities and affective prosody dysfunction in schizophrenia, and indicate that both relatively automatic pre-attentive processes (MMN) and later attention-dependent processes (P300) are involved with accurate auditory emotion identification. These findings provide support for bottom-up (e.g., perceptually based) cognitive remediation approaches.

Published by Elsevier B.V.

1. Introduction

Emotion perception represents a fundamental aspect of social cognition. Disturbances in this basic, low-level, ability lead to deficits in other, high-level areas of social cognition (e.g., mental state attribution). Deficits in perception of emotional cues are well-documented in schizophrenia (Hoekert et al., 2007; Kohler et al., 2010) and lead to downstream impairments in social functioning (Meyer and Kurtz, 2009; Pan et al., 2009; Fiszdon and Johannesen, 2010). In the schizophrenia literature, most studies on emotion perception have used visual stimuli, i.e., still photographs of emotional faces. Much less is known about the mechanisms underlying emotion recognition deficits within the auditory modality.

Emotional prosody or the ability to decode emotions based on the intonation, stress, and rhythm patterns of vocal utterances is critical for normal social interaction and reciprocity (Juslin and Laukka, 2003). Impairments in emotional prosodic processing have been widely reported in schizophrenia (e.g., Shaw et al., 1999; Edwards et al., 2001; Kucharska-Pietura et al., 2005; Bozikas et al., 2006). Whereas substantial evidence suggests that early-stage visual processing contributes to facial affect recognition in schizophrenia (Butler et al., 2009; Norton et al., 2009), little is known about how early auditory processing contributes to affective prosody. It remains unclear whether affective prosody dysfunction in schizophrenia is due to impairment in early auditory sensory processing or later attention-dependent processes. Specialized electroencephalography (EEG) measures within the auditory domain provide a way to conduct a detailed examination of different stages of processing and their relationship to affective prosody.

Schizophrenia patients have deficits in basic auditory perception, such as in pitch processing and tone matching (e.g., Holcomb et al., 1995; Matsumoto et al., 2006). A few reports that examined the relationship between basic auditory processing and prosodic ability in schizophrenia found that deficits in pitch perception affect patients' ability to infer other people's emotions from voice tone (Leitman et al., 2005, 2010a, 2011a; Kantrowitz et al., 2011). Further, Gold et al. (2012) demonstrated that schizophrenia patients are impaired in their ability to identify pitch-based, but not intensity-based, features of emotional stimuli and that these deficits correlate with impaired tone

* Corresponding author at: Mental Illness Research, Education, and Clinical Center (MIRECC), VA Greater Los Angeles Healthcare System, Bldg. 210, Room 117, 11301 Wilshire Blvd., Los Angeles, CA 90073, USA. Tel.: +1 310 478 3711x42953; fax: +1 310 268 4056.
E-mail address: caroljahshan@hotmail.com (C. Jahshan).

0920-9964/$ – see front matter. Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.schres.2012.11.025

matching performance. Recent studies have included neuroscientific methods to address this question. Pinheiro et al. (2012) showed that dysfunctional affective prosody detection in schizophrenia is associated with deficits in very basic level of sensory processing, as indexed by event-related potential (ERP) N100 and P200 amplitudes. Also, fMRI activation during prosodic processing was correlated with a behavioral measure of pitch perception in schizophrenia patients (Leitman et al., 2011b).

The aim of this project was to examine emotion perception in schizophrenia in the auditory domain and its relationships to early (MMN) and late (P300) auditory processing indices. Our primary objective was to examine the strength of the association between each auditory processing stage and affective prosody recognition. Our secondary objective was to assess whether the relationship of auditory to prosodic processing is modality specific by including a facial emotion identification task. We hypothesized that both MMN and P300 will be strongly correlated with auditory, but not with visual, emotion recognition.

2. Method

2.1. Participants

Thirty six patients with schizophrenia and 18 healthy comparison subjects participated in the study. All participants were between the ages of 24 and 65, had an IQ over 70, and were sufficiently fluent in English to understand the procedures. Subjects were excluded if they had substance dependence in the last 6 months, substance abuse in the last month, positive drug toxicology screen, electroconvulsive therapy in the last 6 months, hearing or vision deficits, neurological illness (e.g., seizures, stroke), or significant head injury with loss of consciousness for greater than 15 min.

Patients with schizophrenia were recruited from outpatient treatment clinics at the Veterans Affairs Greater Los Angeles Healthcare System (VAGLAHS) and from board-and-care residences in the community. Diagnosis was based on the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I; First et al., 1997). Patients were considered to be clinically stable, defined as no psychiatric medication changes in the past six weeks, no inpatient hospitalization in the past three months, and no changes in housing in the past two months. All patients were taking antipsychotic medication at the time of testing; 33 were receiving second-generation medications, 1 was taking a first-generation medication, 1 was taking both a first and second-generation medication, and 1 was missing medication information.

Healthy controls were recruited through newspaper and internet advertisements and were screened with the SCID-I and SCID-II (First et al., 1996). They were excluded if they met criteria for current or past psychotic disorder, current Axis I mood disorder, cluster A personality disorder, or if they reported a history of psychosis in a first-degree relative. All participants had the capacity to give informed consent and provided written informed consent after all procedures were explained in accordance with procedures approved by the Institutional Review Boards at the University of California, Los Angeles and VAGLAHS.

2.2. Measures

2.2.1. Clinical and social functioning ratings

Psychiatric symptoms were rated using the Scale for the Assessment of Negative Symptoms (SANS; Andreasen, 1984a) and the Scale for the Assessment of Positive Symptoms (SAPS; Andreasen, 1984b). The 30-item, self-report version of the Specific Levels of Functioning Scale (SLOF; Schneider and Struening, 1983) was used to assess real-world functional performance. We report the domain and total scores for the SANS, SAPS, and SLOF in Table 1.

Table 1
Demographic and clinical characteristics.

	Patients (N = 36) Mean (SD)	Controls (N = 18) Mean (SD)
Age	47.7 (10.0)	45.8 (7.9)
Gender (% male)	69.4%	77.8%
Personal education**	12.6 (1.8)	15.4 (1.2)
Parental education*	12.0 (4.2)	14.8 (3.0)
Race (% white)	52.8%	72.2%
Duration of illness	24.3 (11.5)	–
SANS affective flattening	1.64 (1.50)	–
SANS alogia	0.50 (0.88)	–
SANS avolition	2.61 (0.99)	–
SANS anhedonia	2.56 (1.12)	–
SANS total score	20.07 (10.54)	–
SAPS hallucinations	1.10 (1.58)	–
SAPS delusions	1.72 (1.63)	–
SAPS bizarre behavior	0.52 (0.77)	–
SAPS thought disorder	1.13 (1.41)	–
SAPS total score	9.22 (10.49)	–
SLOF interpersonal relationships	3.71 (0.91)	4.18 (0.60)
SLOF socially acceptable behaviors	4.56 (0.36)	4.70 (0.26)
SLOF independent living skills	4.22 (1.01)	4.94 (0.12)
SLOF work skills	3.96 (0.62)	4.77 (0.36)
SLOF total score	16.33 (2.20)	18.59 (0.91)

SANS = Scale for the Assessment of Negative Symptoms; SAPS = Scale for the Assessment of Positive Symptoms; SLOF = Specific Levels of Functioning Scale; A higher score on the SLOF indicates better functioning.
* $p<0.05$.
** $p<0.001$.

2.2.2. MMN paradigm

MMN was measured using a passive attention auditory oddball paradigm. Subjects were presented with binaural tones (1 kHz 85 dB sound pressure level, with 10 ms rise/fall) with a fixed stimulus onset asynchrony of 500 ms, using E-Prime 2.0. Standard (90% probability; 50 ms duration) and duration-deviant (10% probability; 100 ms duration) tones were presented in a fixed, pseudorandom order (so that at least 4 standard tones were presented between deviant tones) using foam insert earphones. Two-thousand total trials were administered. During the 20-minute EEG recording, subjects were instructed to watch a silent movie to divert attention from the stimuli.

2.2.3. P300 paradigm

Following the MMN, P300 was measured using an active attention auditory oddball paradigm. Subjects were presented with standard (88% probability; 1 kHz frequency; 100 ms duration) and target (12% probability; 1.5 kHz frequency; 100 ms duration) tones with a variable SOA of 1, 2, or 3 s. They were instructed to pay attention to the stimulus stream and press a button every time they heard the high-pitched tone (infrequent target stimulus) while disregarding the low-pitched tone (standard stimulus). There were a total of 600 trials.

2.2.4. EEG recording and analysis

EEG recordings were acquired with a 64-channel BioSemi ActiveTwo amplifier (Biosemi B. V., Amsterdam, Netherlands). Additional electrodes were placed above and below the left eye and at the outer canthi of both eyes to monitor blinks and eye movements. Each active electrode was measured online with respect to a common mode sense electrode during data collection, forming a monopolar channel. Data were sampled at 1024 Hz at DC to 100 Hz. An additional electrode was placed at the nose tip. For MMN, data were re-referenced offline to this electrode. For P300, data were re-referenced offline to the averaged mastoid reference.

Data processing was performed using BrainVision Analyzer 2 (Brain Products, Gilching, Germany). Based on visual inspection, bad electrodes were removed from the recording and a spherical spline interpolation was used to recreate the electrode (Perrin et al., 1989;

Picton et al., 2000). Eyeblinks were removed from the data using a regression-based algorithm (Gratton et al., 1983).

For MMN, data were epoched at −100 to 500 ms relative to stimulus onset and were baseline corrected to the average of the prestimulus interval. Epochs were high-pass filtered at 1 Hz and low-pass filtered at 20 Hz (zero phase shift, 24 dB/octave rolloff) to remove any residual high-frequency artifact. Epochs that contained activity exceeding ±75 μV at frontocentral electrode sites where MMN is commonly seen (AF3, AF4, AF7, AF8, AFz, Cz, F1, F2, F3, F4, F5, F6, F7, F8, FC1, FC2, FC3, FC4, FC5, FC6, FCz, Fp1, Fp2, Fpz, Fz) were automatically rejected. At least 88% of the trials were accepted for all subjects. MMN waveforms were generated by subtracting standard from deviant averaged waveforms. MMN amplitude was measured as the mean voltage in the 150–240 ms latency range.

For P300, continuous data were epoched at −100 to 1000 ms relative to stimulus onset and were baseline corrected to the average of the prestimulus interval. Epochs were high-pass filtered at 0.1 Hz and low-pass filtered at 20 Hz (zero phase shift, 24 dB/octave rolloff). Epochs that contained activity exceeding ±75 μV at electrode sites P1, P3, Pz, P2, and P4 were automatically rejected. At least 95% of the trials were accepted for all subjects. P300 was identified as the largest positive value between 290 and 450 ms at electrode Pz in the target condition.

2.2.5. Tone matching

Pitch perception was measured using a tone matching task (Strous et al., 1995) in which participants were asked to discriminate pitch differences between pairs of tones. Two 100 ms tones, separated by a 500 ms intertone interval, were played in each trial. Participants responded by pressing 1 of the 2 keys to indicate whether the 2 tones were of the same or different pitches. Pairs of tones could range from identical pitch to 50% of an octave apart. There were a total of 94 trials presented in randomized order. Half of the trials (n = 47) were identical. There were 7 trials each for differentials of 50%, 20%, 10%, 5%, and 2.5%, and 6 trials each for differentials of 1.5% and 1%.

2.2.6. Auditory emotion recognition

Recognition of affective prosody was measured using a task (Juslin and Laukka, 2001) in which participants identified emotion for spoken sentences. Stimuli consisted of audio recordings of 4 actors saying semantically neutral sentences with different voice tones. The sentences were either statements or questions (e.g., "It is eleven o'clock", "Is it eleven o'clock?") and conveyed in 5 emotions (fear, anger, happiness, sadness, disgust) or no emotion. The task consisted of 88 trials: 16 trials per emotion as well as 8 neutral trials. After each trial, a list of 6 possible choices was presented and the participant selected one emotion.

2.2.7. Visual emotion recognition

Participants identified facial expressions in photographs from the standardized stimulus set developed by Ekman (2004). The test included photos of 8 different posers displaying facial expressions of 6 emotions (fear, anger, happiness, sadness, disgust, surprise) plus neutral expressions. There were a total of 56 trials, 8 per facial expression. On each trial, a photo and a list of the 7 possible expressions were simultaneously presented on the screen. The photo was displayed until the participant made a key press or for a maximum of 5 s.

2.3. Statistical analyses

Independent samples t-tests and chi-square tests were used to assess group differences for continuous and categorical demographic variables, respectively. Electrodes were averaged together to examine MMN at midline sites (Fpz, AFz, Fz, FCz, Cz), and P300 at parietal sites (P1, P3, Pz, P2, P4). One-sample t-tests were conducted for each ERP component, separately for each group, to determine if activity was significantly different from zero. Independent samples t-tests were conducted to investigate group differences in MMN and P300 (at the averaged electrode sites), tone matching, and social functioning. Two separate repeated measures analysis of variance (rmANOVA) with emotion as the within-subject factor and group as the between-subject factor were conducted to assess group differences in affective prosody and facial emotion recognition. Relationships between the ERP and behavioral measures were investigated using Pearson correlations within the patient group. Lastly, we conducted a multiple regression analysis with MMN and P300 as predictors of affective prosody to examine the unique contributions of these indices to auditory emotion recognition.

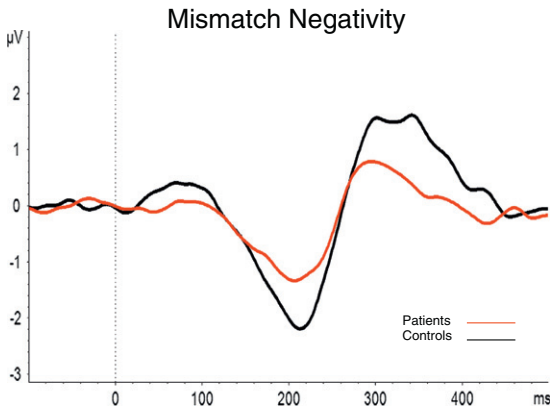

Fig. 1. Grand average MMN waveforms at electrode Fz for healthy controls (in black) and schizophrenia patients (in red).

3. Results

3.1. Demographic and clinical characteristics

Demographic, symptom, and social functioning ratings can be seen in Table 1. There were no significant group differences in age or gender distribution. However, patients had significantly lower personal (t [52] = 5.94, p < 0.001) and parental (t [48] = 2.50, p = 0.02) education than controls. Because parental education can be a reflection of socio-economic status, we examined its possible relationship to the main variables of interest. There were no significant associations between parental education or illness duration and MMN, P300, or affective prosody within the patient group. Patients were relatively chronic and exhibited mild to moderate levels of symptomatology.

3.2. Group differences in ERP measures

One-sample t-tests for MMN and P300 were significant (p's < 0.001) for both groups. Patients had reduced MMN amplitude relative to healthy controls for the averaged midline sites, −0.93 (0.69) and −1.39 (0.78) μV, respectively (t [51] = −2.19, p = 0.03). Fig. 1 shows a grand average MMN waveform at Fz for each group. Patients' P300 amplitude was reduced relative to controls at the averaged parietal sites, 6.19 (4.16) and 9.02 (5.81) μV, respectively (t [52] = 2.06, p = 0.04). Fig. 2 shows a grand average P300 waveform for the standard and target conditions at Pz for each group. Topographical maps of MMN and P300 are shown in Fig. 3 for patients and controls separately. MMN was largest at frontocentral electrodes with the expected polarity inversion at temporo-parietal electrodes. P300 was maximal at parietal electrodes.

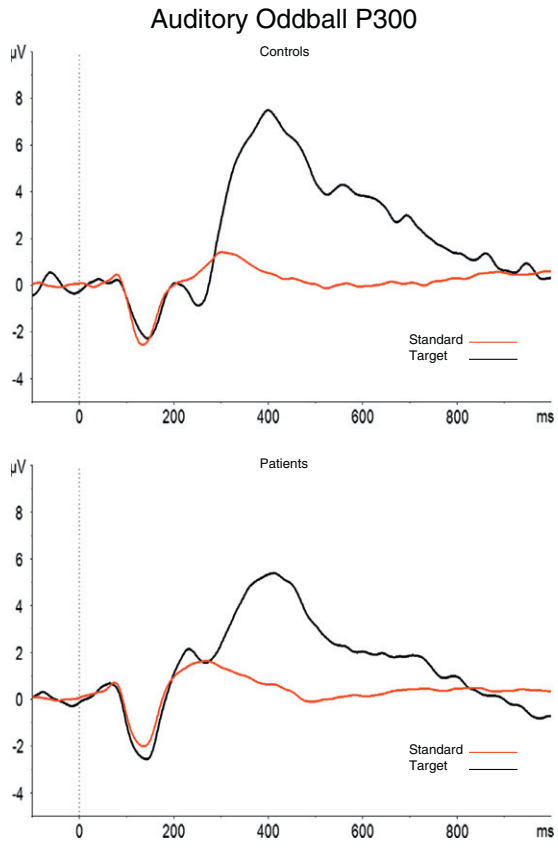

Fig. 2. Grand average P300 waveforms for the standard and target conditions at electrode Pz for healthy controls (top panel) and schizophrenia patients (bottom panel).

3.3. Group differences in behavioral measures

Table 2 displays the participants' performance on the tone matching task. Both groups' overall performance was significantly different from chance performance (p's<0.001) based on the binomial distribution. The total accuracy score for patients was significantly lower than controls (t [43] $= 2.60$, $p = 0.01$). Patients were significantly less accurate than controls in discriminating between pairs of tones that differed by 1% (t [43] $= 3.89$, $p<0.001$) and 2.5% (t [43] $= 2.63$, $p = 0.01$) in pitch, but did not differ significantly at the other pitch differentials.

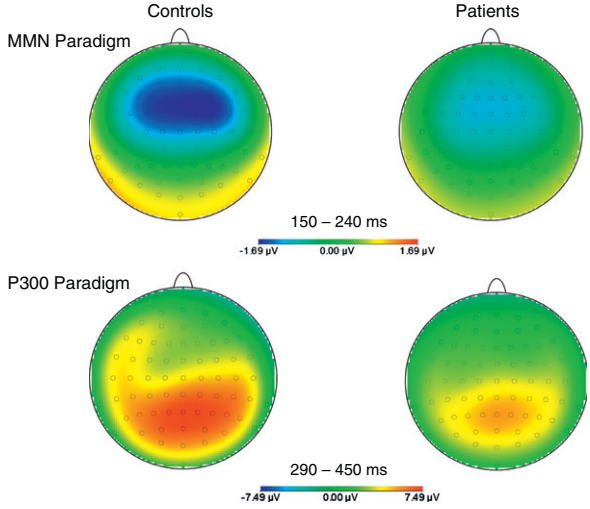

Fig. 3. Topographical maps of MMN and P300 for healthy controls (left panel) and schizophrenia patients (right panel).

Table 2
Tone matching performance.

Pitch differentials	Patients (N = 31) Mean (SD)	Controls (N = 14) Mean (SD)
1%**	0.77 (1.02)	2.36 (1.69)
1.5%	1.32 (1.33)	2.07 (2.16)
2.5%*	1.87 (1.67)	3.50 (2.41)
5%	4.00 (2.50)	4.79 (2.45)
10%	5.77 (1.61)	5.71 (1.59)
20%	6.42 (0.89)	6.57 (0.85)
50%	6.65 (1.02)	6.86 (0.36)
0%	44.48 (3.91)	46.07 (1.49)
Total score*	71.29 (7.50)	77.93 (8.82)

Highest possible score is 6 for the 1% and 1.5% differentials, 7 for the 2.5% to 50% differentials, 47 for the 0% differential, and 94 for the total score.
* $p<0.05$.
** $p<0.001$.

The rmANOVA for the affective prosody task revealed significant main effects of group (F [1, 46] $= 14.91$, $p<0.001$) and emotion (F [5, 230] $= 7.57$, $p<0.001$) but no significant emotion × group interaction. Patients were significantly less accurate than controls in identifying the correct emotion from voice tone. The main effect of emotion was due to better accuracy across groups for neutral, anger, and sadness, and poorer for disgust (see Table 3).

The rmANOVA for the facial emotion identification task revealed significant main effects of group (F [1, 50] $= 9.73$, $p = 0.003$) and emotion (F [6, 300] $= 18.32$, $p<0.001$) but no significant emotion × group interaction. Patients were significantly less accurate than controls in identifying the correct emotion from facial expression. The main effect of emotion was due to better accuracy across groups for happiness, surprise, and neutral, and poorer for fear.

3.4. Correlations with ERPs within the patient group

For this analysis, MMN and P300 (at the averaged electrodes), the total scores on the tone matching, affective prosody, and facial emotion tasks, as well as the total SANS, SAPS, and SLOF scores were included. There were significant correlations between the affective prosody task and both MMN ($r = -0.46$; $p = 0.007$) and P300 ($r = 0.51$; $p = 0.002$) but no significant correlations between the facial emotion task and either ERP measure. We found no significant associations between the emotion perception measures and SANS, SAPS, or SLOF. The tone matching task was not correlated with any of the

Table 3
Emotion recognition performance.

	Affective Prosody Task		Facial Emotion Identification Task	
	Patients (N = 34) Mean (SD)	Controls (N = 14) Mean (SD)	Patients (N = 34) Mean (SD)	Controls (N = 18) Mean (SD)
Fear	4.09 (2.70)	5.79 (1.67)	3.35 (2.33)	5.56 (2.48)
Anger	5.85 (2.06)	7.43 (2.34)	5.53 (2.09)	6.56 (1.65)
Happiness	4.62 (2.86)	5.36 (2.47)	7.47 (1.50)	7.78 (0.65)
Sadness	5.62 (2.76)	7.21 (2.83)	5.59 (2.02)	6.28 (2.14)
Disgust	3.79 (2.25)	4.43 (2.10)	5.88 (2.38)	7.11 (1.32)
Surprise	–	–	6.76 (1.86)	7.33 (0.77)
Neutral	5.24 (2.17)	7.21 (0.97)	6.71 (1.49)	7.28 (0.96)
Total Score	29.21 (6.94)	37.43 (6.06)**	41.29 (8.01)	47.89 (5.50)*

Affective Prosody Task: Highest possible score is 16 for each emotion (8 for neutral) and 88 for the total score. Facial Emotion Identification Task: Highest possible score is 8 for each emotion and 56 for the total score.
* $p<0.05$.
** $p<0.001$.

Table 4
Correlations between ERP and behavioral measures in schizophrenia patients.

	MMN	P300	TMT	APT	FEIT	SLOF	SANS	SAPS
MMN	–							
P300	−0.06	–						
TMT	−0.08	0.20	–					
APT	−0.46**	0.51**	0.26	–				
FEIT	−0.23	0.17	0.07	0.44**	–			
SLOF	−0.05	−0.01	0.37	−0.04	0.11	–		
SANS	0.09	−0.07	0.07	−0.32	−0.07	−0.14	–	
SAPS	0.28	0.07	−0.12	0.07	−0.16	−0.31	0.06	–

MMN = Mismatch Negativity; TMT = Tone Matching Task; APT = Affective Prosody Task; FEIT = Facial Emotion Identification Task; SLOF = Specific Levels of Functioning Scale; SANS = Scale for the Assessment of Negative Symptoms; SAPS = Scale for the Assessment of Positive Symptoms.
** $p<0.01$; two-tailed.

measures, and MMN/P300 were not correlated with the SANS, SAPS, or SLOF (Table 4).

A linear regression with MMN and P300 entered simultaneously as independent variables and the affective prosody task total score as the dependent variable revealed significant regression coefficients for both factors (MMN: $\beta = -0.45$; $p = 0.002$; P300: $\beta = 0.52$; $p < 0.001$). The two-predictor model accounted for 49% of the variance in affective prosody, (F [2, 30] = 14.18, $p < 0.001$).

4. Discussion

We explored links between different stages of auditory processing, indexed by MMN and P300, and affective prosody detection in patients with schizophrenia. Relative to controls, patients had significantly reduced MMN and P300 amplitudes, showed impaired recognition of emotions whether conveyed by facial expressions or speech, and performed poorly on the tone matching task. MMN and P300 significantly correlated with affective prosody within the patient group. As hypothesized, these relationships were modality specific: MMN and P300 did not correlate with visual emotion recognition. Regression analysis revealed that MMN and P300 accounted for 49% of the variance in the affective prosody task. Our results support previous suggestions of an association between basic auditory processing abnormalities and affective prosody dysfunction in schizophrenia (Leitman et al., 2010a, 2010b; Pinheiro et al., 2012). By examining the time course of auditory processing with a high level of precision, we demonstrated that both early (MMN) and late (P300) stages of auditory processing contributed roughly equally to affective prosody detection in patients with schizophrenia.

We are not aware of any other published studies that examined early and late auditory neural correlates of dysfunctional affective prosody in schizophrenia. Our general conclusions of the importance of early auditory processing for prosody detection in schizophrenia are consistent with the conclusions from Javitt et al. (2000) and Leitman et al. (2010a). In contrast to these studies, we failed to detect an association between tone matching and both affective prosody and MMN, perhaps due to some differences in methods. Unlike several previous studies that used pitch deviance as the basis of the MMN (e.g., Javitt et al., 2000; Salisbury et al., 2002; Leitman et al., 2011a), we used a duration-deviant version. We selected this type of MMN because it may be a more sensitive index of auditory system dysfunction in schizophrenia than pitch MMN (e.g., Michie, 2001; Umbricht and Krljes, 2005). Despite this difference, we still found a significant relationship between MMN and pitch-based affective prosody items, suggesting that a pervasive preattentive auditory impairment underlies affective prosody dysfunction in schizophrenia. However, our use of duration MMN with pitch-oddball P300 could explain the lack of an association between these two ERPs. Leitman et al. (2010b), using pitch deviants for both MMN and P300, did find these components correlated in schizophrenia.

The study has several limitations. First, our groups were not matched on parental education. However, this variable was not correlated with our primary outcome variables (MMN, P300, affective prosody) and we were mainly interested in relationships between affective prosody and auditory processing within the patient group. Second, our sample consisted of chronic patients who were receiving antipsychotic medications at the time of testing, which raises the question of whether our results generalize to a recent-onset or unmedicated sample. Our patients showed substantial MMN and P300 amplitude reductions despite the suggestion that some second-generation antipsychotics may improve these ERPs (Korostenskaja and Kahkonen, 2009; Horton et al., 2011). Moreover, chronicity did not account for our findings as there were no significant associations between illness duration and ERP or behavioral measures. Third, we did not examine patterns of correlations within healthy controls because we had fewer controls than patients. Fourth, the sample size of our patient group was relatively small, which could have contributed to our failure to replicate the previously reported relationship between tone matching and affective prosody (e.g., Kantrowitz et al., 2011; Gold et al., 2012).

In summary, our results are consistent with a cascade model in which perceptual processes lead to social cognitive abilities (e.g., Javitt, 2009). Early and late auditory ERPs independently contributed to the ability to recognize emotions from voice tone. Hence, both relatively automatic pre-attentive processes (i.e., MMN) and later attention-dependent processes (i.e., P300) appear to be needed for accurate auditory emotion identification. Bottom-up cognitive remediation strategies that target basic auditory processes (e.g., Adcock et al., 2009; Dale et al., 2010) would be expected to lead to improvements in this important aspect of social cognition.

Role of funding source

Funding for this study was provided by grants from the Department of Veterans Affairs Sierra Pacific Mental Illness Research, Education and Clinical Center (MIRECC) and the National Institute of Mental Health (MH065707 to MFG). The MIRECC and NIH had no further role in study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

Contributors

Drs. Jahshan and Green designed the study. Dr. Wynn supervised the data analysis conducted by Dr. Jahshan. Dr. Jahshan took the lead on writing the manuscript with the help in writing from Drs. Green and Wynn. All authors contributed to and have approved the final manuscript.

Conflict of interest

All authors declare that they have no conflicts of interest arising from this manuscript.

Acknowledgments

Writing of this manuscript was supported by the Department of Veterans Affairs Office of Academic Affiliations Advanced Fellowship Program in Mental Illness Research and Treatment. The authors would like to thank Crystal Gibson, Cory Tripp, Katie Weiner, Mark McGee, Christen Waldon and Amanda Bender for their assistance with recruitment and testing. The authors would also like to thank Petri Laukka, Ph.D. and David I. Leitman, Ph.D. for providing the affective prosody task and tone matching task, respectively.

References

Adcock, R.A., Dale, C., Fisher, M., Aldebot, S., Genevsky, A., Simpson, G.V., Nagarajan, S., Vinogradov, S., 2009. When top-down meets bottom-up: auditory training enhances verbal memory in schizophrenia. Schizophr. Bull. 35 (6), 1132–1141.
Andreasen, N.C., 1984a. The Scale for the Assessment of Negative Symptoms (SANS). The University of Iowa, Iowa City, IA.
Andreasen, N.C., 1984b. The Scale for The assessment of Positive Symptoms (SAPS). The University of Iowa, Iowa City, IA.
Bozikas, V.P., Kosmidis, M.H., Anezoulaki, D., Giannakou, M., Andreou, C., Karavatos, A., 2006. Impaired perception of affective prosody in schizophrenia. J. Neuropsychiatry Clin. Neurosci. 18 (1), 81–85.
Butler, P.D., Abeles, I.Y., Weiskopf, N.G., Tambini, A., Jalbrzikowski, M., Legatt, M.E., Zemon, V., Loughead, J., Gur, R.C., Javitt, D.C., 2009. Sensory contributions to impaired emotion processing in schizophrenia. Schizophr. Bull. 35 (6), 1095–1107.

Dale, C.L., Findlay, A.M., Adcock, R.A., Vertinski, M., Fisher, M., Genevsky, A., Aldebot, S., Subramaniam, K., Luks, T.L., Simpson, G.V., Nagarajan, S.S., Vinogradov, S., 2010. Timing is everything: neural response dynamics during syllable processing and its relation to higher-order cognition in schizophrenia and healthy comparison subjects. Int. J. Psychophysiol. 75 (2), 183–193.
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.
Ekman, P., 2004. Subtle Expression Training Tool (SETT) & Micro Expression Training Tool (METT) [computer program]. (www.paulekman.com).
First, M.B., Gibbon, M., Spitzer, R.L., Williams, J.B.W., Benjamin, L., 1996. Structured Clinical Interview for DSM-IV Avis II Personality Disorders. Biometrics Research Department, New York State Psychiatric Institute, New York, NY.
First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.W., 1997. Structured Clinical Interview for DSM-IV Axis I Disorders — Patient Edition. Biometrics Research Department, New York State Psychiatric Institute, New York, NY.
Fiszdon, J.M., Johannesen, J.K., 2010. Functional significance of preserved affect recognition in schizophrenia. Psychiatry Res. 176 (2–3), 120–125.
Gold, R., Butler, P., Revheim, N., Leitman, D.I., Hansen, J.A., Gur, R.C., Kantrowitz, J.T., Laukka, P., Juslin, P.N., Silipo, G.S., Javitt, D.C., 2012. Auditory emotion recognition impairments in schizophrenia: relationship to acoustic features and cognition. Am. J. Psychiatry 169 (4), 424–432.
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.
Hoekert, M., Kahn, R.S., Pijnenborg, M., Aleman, A., 2007. Impaired recognition and expression of emotional prosody in schizophrenia: review and meta-analysis. Schizophr. Res. 96 (1–3), 135–145.
Holcomb, H.H., Ritzl, E.K., Medoff, D.R., Nevitt, J., Gordon, B., Tamminga, C.A., 1995. Tone discrimination performance in schizophrenic patients and normal volunteers: impact of stimulus presentation levels and frequency differences. Psychiatry Res. 57 (1), 75–82.
Horton, J., Millar, A., Labelle, A., Knott, V.J., 2011. MMN responsivity to manipulations of frequency and duration deviants in chronic, clozapine-treated schizophrenia patients. Schizophr. Res. 126 (1–3), 202–211.
Javitt, D.C., 2009. When doors of perception close: bottom-up models of disrupted cognition in schizophrenia. Annu. Rev. Clin. Psychol. 5, 249–275.
Javitt, D.C., Shelley, A., Ritter, W., 2000. Associated deficits in mismatch negativity generation and tone matching in schizophrenia. Clin. Neurophysiol. 111 (10), 1733–1737.
Juslin, P.N., Laukka, P., 2001. Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. Emotion 1 (4), 381–412.
Juslin, P.N., Laukka, P., 2003. Communication of emotions in vocal expression and music performance: different channels, same code? Psychol. Bull. 129 (5), 770–814.
Kantrowitz, J.T., Leitman, D.I., Lehrfeld, J.M., Laukka, P., Juslin, P.N., Butler, P.D., Silipo, G., Javitt, D.C., 2011. Reduction in Tonal Discriminations Predicts Receptive Emotion Processing Deficits in Schizophrenia and Schizoaffective Disorder. Schizophr. Bull. (Epub ahead of print).
Kohler, C.G., Walker, J.B., Martin, E.A., Healey, K.M., Moberg, P.J., 2010. Facial emotion perception in schizophrenia: a meta-analytic review. Schizophr. Bull. 36 (5), 1009–1019.
Korostenskaja, M., Kahkonen, S., 2009. What do ERPs and ERFs reveal about the effect of antipsychotic treatment on cognition in schizophrenia? Curr. Pharm. Des. 15 (22), 2573–2593.
Kucharska-Pietura, K., David, A.S., Masiak, M., Phillips, M.L., 2005. Perception of facial and vocal affect by people with schizophrenia in early and late stages of illness. Br. J. Psychiatry 187, 523–528.
Leitman, D.I., Foxe, J.J., Butler, P.D., Saperstein, A., Revheim, N., Javitt, D.C., 2005. Sensory contributions to impaired prosodic processing in schizophrenia. Biol. Psychiatry 58 (1), 56–61.
Leitman, D.I., Laukka, P., Juslin, P.N., Saccente, E., Butler, P., Javitt, D.C., 2010a. Getting the cue: sensory contributions to auditory emotion recognition impairments in schizophrenia. Schizophr. Bull. 36 (3), 545–556.
Leitman, D.I., Sehatpour, P., Higgins, B.A., Foxe, J.J., Silipo, G., Javitt, D.C., 2010b. Sensory deficits and distributed hierarchical dysfunction in schizophrenia. Am. J. Psychiatry 167 (7), 818–827.
Leitman, D.I., Sehatpour, P., Garidis, C., Gomez-Ramirez, M., Javitt, D.C., 2011a. Preliminary evidence of pre-attentive distinctions of frequency-modulated tones that convey affect. Front. Hum. Neurosci. 5, 96.
Leitman, D.I., Wolf, D.H., Laukka, P., Ragland, J.D., Valdez, J.N., Turetsky, B.I., Gur, R.E., Gur, R.C., 2011b. Not pitch perfect: sensory contributions to affective communication impairment in schizophrenia. Biol. Psychiatry 70 (7), 611–618.
Matsumoto, K., Samson, G.T., O'Daly, O.D., Tracy, D.K., Patel, A.D., Shergill, S.S., 2006. Prosodic discrimination in patients with schizophrenia. Br. J. Psychiatry 189, 180–181.
Meyer, M.B., Kurtz, M.M., 2009. Elementary neurocognitive function, facial affect recognition and social-skills in schizophrenia. Schizophr. Res. 110 (1–3), 173–179.
Michie, P.T., 2001. What has MMN revealed about the auditory system in schizophrenia? Int. J. Psychophysiol. 42 (2), 177–194.
Norton, D., McBain, R., Holt, D.J., Ongur, D., Chen, Y., 2009. Association of impaired facial affect recognition with basic facial and visual processing deficits in schizophrenia. Biol. Psychiatry 65 (12), 1094–1098.
Pan, Y.J., Chen, S.H., Chen, W.J., Liu, S.K., 2009. Affect recognition as an independent social function determinant in schizophrenia. Compr. Psychiatry 50 (5), 443–452.
Perrin, F., Pernier, J., Bertrand, O., Echallier, J.F., 1989. Spherical splines for scalp potential and current density mapping. Electroencephalogr. Clin. Neurophysiol. 72 (2), 184–187.
Picton, T.W., Bentin, S., Berg, P., Donchin, E., Hillyard, S.A., Johnson Jr., R., Miller, G.A., Ritter, W., Ruchkin, D.S., Rugg, M.D., Taylor, M.J., 2000. Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. Psychophysiology 37 (2), 127–152.
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.
Salisbury, D.F., Shenton, M.E., Griggs, C.B., Bonner-Jackson, A., McCarley, R.W., 2002. Mismatch negativity in chronic schizophrenia and first-episode schizophrenia. Arch. Gen. Psychiatry 59 (8), 686–694.
Schneider, L.C., Struening, E.L., 1983. SLOF: a behavioral rating scale for assessing the mentally ill. Soc. Work Res. Abstr. 19 (3), 9–21.
Shaw, R.J., Dong, M., Lim, K.O., Faustman, W.O., Pouget, E.R., Alpert, M., 1999. The relationship between affect expression and affect recognition in schizophrenia. Schizophr. Res. 37 (3), 245–250.
Strous, R.D., Cowan, N., Ritter, W., Javitt, D.C., 1995. Auditory sensory ("echoic") memory dysfunction in schizophrenia. Am. J. Psychiatry 152 (10), 1517–1519.
Umbricht, D., Krljes, S., 2005. Mismatch negativity in schizophrenia: a meta-analysis. Schizophr. Res. 76 (1), 1–23.