Basic auditory processing deficits and their association with auditory emotion recognition in schizophrenia☆

Michael S. Kraus [a,*], Trina M. Walker [a], L. Fredrik Jarskog [b], Robert A. Millet [c], Richard S.E. Keefe [a]

[a] Department of Psychiatry & Behavioral Sciences, Duke University Medical Center, 10 Duke Medicine Circle, Durham, NC 27710, USA
[b] North Carolina Psychiatric Research Center, Department of Psychiatry, University of North Carolina at Chapel Hill, 101 Manning Dr # 1, Chapel Hill, NC 27514, USA
[c] Carolina Behavioral Care, 4102 Ben Franklin Blvd Durham, NC 27704, USA

ARTICLE INFO

ABSTRACT

Background: Individuals with schizophrenia are impaired in their ability to recognize emotions based on vocal cues and these impairments are associated with poor global outcome. Basic perceptual processes, such as auditory pitch processing, are impaired in schizophrenia and contribute to difficulty identifying emotions. However, previous work has focused on a relatively narrow assessment of auditory deficits and their relation to emotion recognition impairment in schizophrenia.
Methods: We have assessed 87 patients with schizophrenia and 73 healthy controls on a comprehensive battery of tasks spanning the five empirically derived domains of auditory function. We also explored the relationship between basic auditory processing and auditory emotion recognition within the patient group using correlational analysis.
Results: Patients exhibited widespread auditory impairments across multiple domains of auditory function, with mostly medium effect sizes. Performance on all of the basic auditory tests correlated with auditory emotion recognition at the p < .01 level in the patient group, with 9 out of 13 tests correlating with emotion recognition at r = 0.40 or greater. After controlling for cognition, many of the largest correlations involved spectral processing within the phase-locking range and discrimination of vocally based stimuli.
Conclusions: While many auditory skills contribute to this impairment, deficient formant discrimination appears to be a key skill contributing to impaired emotion recognition as this was the only basic auditory skill to enter a step-wise multiple regression after first entering a measure of cognitive impairment, and formant discrimination accounted for significant unique variance in emotion recognition performance after accounting for deficits in pitch processing.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

Schizophrenia is a debilitating disorder in which patients often struggle to live independently, maintain relationships and maintain employment (Bowie et al., 2006). Antipsychotic medications can improve positive symptoms of psychosis but often leave functioning impaired. Impaired processing of social information is a key determinant of functional outcomes and is thus a promising target for improving functioning in patients with schizophrenia (Brekke et al., 2005; Cornblatt et al., 2012).

One key component of social cognition is the ability to recognize emotions in other people. While children may be quite explicit about their emotions, adolescents and adults tend to be much more subtle. As one matures, deciphering subtle facial and vocal cues of emotion becomes a key skill for navigating personal and professional relationships. Individuals with schizophrenia are impaired in their ability to recognize emotions based on facial and vocal cues and these impairments are associated with poor global outcome (Brekke et al., 2005). Preliminary findings have suggested that emotion recognition training can improve social cognition in patients with schizophrenia (Sachs et al., 2012). However, these training regimens have largely focused on recognition of emotion in facial expressions and have left training of vocal emotion recognition largely untapped. Yet patients with schizophrenia struggle to understand emotion from vocal tone and this difficulty impacts their lives.

Emotion is conveyed in several separable vocal characteristics of the speaker including level and contours of vocal pitch, loudness, and tempo/rhythm, as well as vowel formant structure (amplitude peaks in the frequency spectrum of the speaker's voice) (Banse and Scherer, 1996). Patients have difficulty recognizing emotion from auditory cues

embedded in a speaker's vocal characteristics (Corcoran et al., 2015; Cornblatt et al., 2012; Dickey et al., 2010; Dickey et al., 2008; Leitman et al., 2007; Leitman et al., 2010; Niemi et al., 2005; Thompson et al., 2011). A meta-analysis concluded that the effect size of auditory emotion recognition deficits in schizophrenia is large (Hoekert et al., 2007) and deficits are evident early in disease progression (Corcoran et al., 2015; Grisham et al., 2011; Thompson et al., 2011). In turn, deficits in auditory emotion processing have been linked to auditory hallucinations (Badcock, 2010; Rossell and Boundy, 2005; Shea et al., 2007) and impaired global outcome (Leitman et al., 2005) and are thus an attractive target for intervention. Indeed, early results suggest that training of basic auditory skills can have far reaching benefits for basic auditory processing, ability to follow speech in noise, and improved verbal long term and working memory (Chandrasekaran et al., 2009; Franklin et al., 2008; Kraus and Chandrasekaran, 2010; Nikjeh et al., 2009). Additionally, current auditory cognitive remediation approaches normalize disturbed M100 signals during phoneme discrimination, suggesting some effect on speech perception (Adcock et al., 2009). Auditory training could be designed to improve vocal emotion recognition in individuals with schizophrenia and those at high-risk of developing psychosis. However, current auditory cognitive remediation approaches emphasize training of temporal auditory processing, which may not be a key auditory skill in recognizing emotion in voice.

Individuals with schizophrenia exhibit deficits in pitch discrimination of simple sine waves (Holcomb et al., 1995; Javitt et al., 1997; March et al., 1999) and MMN to frequency modulated tones (Kantrowitz et al., 2015). These frequency processing deficits are associated with impaired vocal emotion recognition (Kantrowitz et al., 2015; Leitman et al., 2005; Leitman et al., 2010), suggesting that a remediation approach that targets frequency processing may improve emotion recognition. Patients with schizophrenia are also impaired in their ability to recognize differences in tone duration and volume (Javitt, 2009). However, a comprehensive assessment spanning all domains of auditory function has not been completed in a schizophrenia sample. Of the five empirically defined domains of auditory acuity, including pitch and time, loudness and duration, amplitude modulation, familiar sounds, and general auditory ability (known as "auditory g") (Kidd et al., 2007), only the relationship between measures of pitch processing and emotion processing has been investigated in schizophrenia. Thus it is not clear which vocal signals of emotion individuals with schizophrenia might be missing and what auditory skills should be targeted to improve their ability to recognize emotions in voice.

Individuals with schizophrenia have exhibited pitch discrimination deficits at base frequencies ranging from 500 Hz to 2000 Hz (Holcomb et al., 1995; Javitt et al., 1997; March et al., 1999). At these frequencies, neurons within the auditory system are able to phase-lock their firing to periodic components of the stimulus, allowing for a more precise encoding of stimulus frequency than occurs with a strictly place-coded signal (i.e. which neurons are firing) (Rose et al., 1967). Phase-locking to auditory stimuli is dependent on intact GABAergic modulation (Fukui et al., 2010). Thus a demonstration of impaired phase locking in schizophrenia would potentially suggest a specific pharmacological target. Although previously reported pitch processing deficits in schizophrenia are consistent with impaired phase-locking of auditory neurons, to our knowledge, no previous study has included control conditions to assess discrimination of stimuli outside the phase-locking range. To address this, we have included tests that employ stimuli within the phase-locking range (frequency discrimination 1000 Hz, frequency modulation (FM) 2 Hz, sinusoidal amplitude modulation 60 Hz) as well as control tests that employ stimuli outside the phase-locking range (frequency discrimination 6000 Hz, FM 10 Hz).

While the mean pitch and the pitch contour are important signals of a speaker's emotion, the human voice is much more spectrally complex than simple sine waves, and in that spectral complexity lie further cues to a speaker's emotion. The distinctive resonances of an individual's vocal tract produces speech formants (amplitude peaks in the frequency spectrum) that not only define vowel sounds, but convey the speaker's emotion through shifts in formant frequencies and widths compared to emotionally neutral speech. Healthy individuals can identify emotion valence after very brief exposure to a voiced vowel, suggesting that the ability to distinguish slight differences in formant structure might be an important auditory skill for deciphering the emotion of a speaker (Waaramaa et al., 2010). Individuals with schizophrenia are impaired in categorizing ambiguous phonemes along the "Ba"–"Da" continuum that differ only in the direction of movement of their second formant (Cienfuegos et al., 1999). We therefore included a formant discrimination task in our basic auditory battery and hypothesized that patients with schizophrenia would exhibit impairment in formant discrimination that would be a key contributor to difficulty understanding emotion in voice.

In this study, we have characterized the basic auditory processing ability of schizophrenia subjects on a battery of tasks spanning the five empirically derived domains of auditory function and have explored the relationship between these basic perceptual measures and emotion recognition.

2. Methods and materials

2.1. Subjects

The study was approved by the Duke University institutional review board. Informed consent was obtained from 87 outpatients with DSM-IV schizophrenia or schizoaffective disorder and 73 healthy controls following full explanation of procedures. All subjects were between 18 and 60 years old, and had intact hearing as demonstrated by audiometric thresholds at or below 20 dB at 500 Hz, 1000 Hz, 2000 Hz, 3000 Hz or 4000 Hz. All subjects were screened to rule out a history of significant head injury, neurological or medical conditions that could interfere with interpretation of cognitive/perception data, or substance abuse within the last month/substance dependence with the last 6 months (excluding nicotine). Healthy controls were cognitively normal as determined by the investigator and were screened to rule out personal history of serious mental illness. All patients were maintained on a stable dose of antipsychotic for at least 2 months prior to participation.

2.2. Basic auditory battery

All of the auditory tasks were structured similarly, with two identical stimuli and one deviant to be detected.

2.2.1. Test of basic auditory capabilities (TBAC)

The TBAC is composed of 8 tests of auditory acuity chosen for their favorable psychometric properties and relatively specific factor loadings onto empirically determined auditory processing domains. The first three tests, *Pitch Discrimination*, *Intensity Discrimination* and *Duration Discrimination*, employ a 1 kHz, 250 ms tone as the standard and the target stimulus is varied along the dimension of interest by equal logarithmic steps. *Pulse Train* tests rhythmic change detection ability. *Embedded Tone* tests the ability to detect the presence of the fifth tone (duration varied to manipulate difficulty) in a nine tone sequence, which is absent in the standard. *Temporal Order* tests the ability to discriminate the order in which two tones are presented. *Syllable Sequence* is a speech analog to the temporal order subtest. *Nonword Recognition* assesses the ability to identify a nonsense syllable in noise.

2.2.2. Formant discrimination task

With consultation from Diane Kewley-Port, we adapted the formant discrimination task from Liu and Kewley-Port (2004). Our modified version of the test assessed subjects' ability to distinguish first and second formant frequency shifts in two American English vowels ε ("eh") and Λ ("uh").

2.2.3. Sinusoidal amplitude modulation (SAM)

Although SAM depth was determined to be a discrete domain of auditory ability, it was left out of the original version of the TBAC to limit testing time. However, subjects with schizophrenia have impaired physiological responses to SAM stimuli indicative of impaired phase locking due to reduced GABAergic modulation of pyramidal cell activity in auditory cortex (Teale et al., 2008). Phase locking is crucial to accurate encoding of fundamental frequency as well as speech formants (Johnson et al., 2008; Johnson et al., 2005). We therefore added a test of SAM detection to allow us to assess this crucial discrete domain of auditory function.

2.2.4. Pitch discrimination 6000 Hz

The frequency discrimination task that is included in the TBAC uses a standard tone of 1000 Hz – a frequency at which phase locking plays a significant role in frequency processing (Sek and Moore, 1995). As a control condition, we included another frequency discrimination task with a standard frequency of 6000 Hz – at which phase locking plays very little role in frequency processing compared to place coding signals.

2.2.5. Frequency modulation (FM)

All three tones had the same carrier wave frequency and differed only in that one tone was frequency modulated (depth manipulated to determine threshold). This experiment was run in separate blocks employing frequency modulations of 2 Hz (at which phase-locking plays a significant role in FM detection) and 10 Hz (at which FM detection depends primarily on place signaling).

2.3. Measures of auditory emotion processing

We assessed the ability to recognize emotion based on vocal affective cues in speech and non-speech affective bursts using two different tests.

2.3.1. Comprehensive Affect Testing System (CATS)

We chose this battery because the American English accent of the actor seemed most appropriate for our demographic. The subtests included: emotion prosody discrimination, emotion prosody identification, and the identification of emotion from prosody in the presence of conflicting semantic meaning. In all tests, a male actor portrayed a specific emotion. Subjects were instructed to make their assessment of the emotion based solely on the sound of the actor's voice (Schaffer et al., 2006).

2.3.2. Montreal Affective Voices

The Montreal Affective Voices are freely available stimuli consisting of multiple male and female actors portraying various emotions through affective bursts (e.g. chuckles, groans, sighs etc.). We developed a test of auditory emotion recognition in which subjects identified affective bursts using a 5-alternative forced-choice procedure (choosing from: anger, disgust, fear, happiness, and sadness).

2.4. Measures of cognition

We assessed current cognition with the Brief Assessment of Cognition in Schizophrenia (BACS), a 30-minute battery of tests measuring verbal memory, processing speed, working memory, executive functioning and verbal fluency (Keefe et al., 2004).

2.5. Assessment of symptoms and functioning

We assessed clinical symptoms with the Positive and Negative Syndrome Scale (PANSS) (Kay et al., 1987). For all subjects, we assessed functional capacity with the UCSD Performance-based Skills Assessment (UPSA) (Green et al., 2011).

2.6. Order of assessments

To minimize order effects, we divided the performance based assessments into 2 blocks and counterbalanced the order of assessment separately in patients and controls.

2.7. Data analysis

An emotion composite was constructed by first calculating Z-scores for the CATS Prosody and MAV outcome variables, and then calculating T-scores for the sum of these Z-scores. We used descriptive statistics to generate plots to examine the distribution of each putative deficit in each of the samples, and test for normality. SAM 60, pitch discrimination 6000 Hz, FM 10 Hz and FM 2 Hz were all found to be significantly skewed and each was normalized by natural log transformation. Group differences were assessed by between-groups t-tests and computation of Cohen's d. All correlations were performed in the patient group. Correlations were computed as Pearson's correlation coefficients, and partial correlations, controlling for general performance deficits, were performed when appropriate.

3. Results

3.1. Demographics

The 87 patients and 73 healthy controls did not differ significantly on demographic variables, other than years of education (Table 1).

3.2. Emotion recognition, cognition and functional capacity

Patients with schizophrenia exhibited similar levels of impairment on both tests of emotion recognition (Table 2). The effect size for a composite score combining the 2 batteries was d = 0.60 (a medium effect size). The patients exhibited deficits of similar effect sizes in cognition (d = 0.65) and functional capacity (d = 0.47).

3.3. Basic auditory processing

Factor analysis of the basic auditory tests indicated a single factor solution and thus all analyses were performed at the level of the individual tests. Patients demonstrated widespread impairment of basic auditory processing. Of the 13 tests included in our battery, patients exhibited deficits on 10 (Table 2). Intensity discrimination yielded the largest effect size (d = 0.66) while formant discrimination yielded an effect size of d = 0.53.

Table 1
Demographics.

	Schizophrenia			Controls			t	p
	N	Mean	S.D.	N	Mean	S.D.		
Age	87	41.83	9.60	73	39.49	11.47	1.38	.17
Education (years)	87	12.91	2.62	73	14.52	2.22	−4.22	<.001
Maternal education	74	13.23	2.84	71	13.54	2.9	−0.64	.52
Paternal education	68	12.93	4.15	67	13.76	3.31	−1.29	.20
Gender								
Male	55	63.22%		44	60.27%			
Female	32	36.78%		29	39.73%			
Race								
White	35	40.23%		29	39.73%			
African-American	51	58.62%		41	56.16%			
Other	1	1.15%		3	4.11%			

Table 2
Group means on emotion recognition, cognition, functional capacity, and basic auditory skills measures. CATS: Comprehensive Affect Testing System Prosody Scale, MAV: Montreal Affective Voices, emotion composite: composite score of CATS and MAV, BACS composite: Brief Assessment of Cognition in Schizophrenia composite score, UPSA: UCSD Performance-based Skills Assessment, SAM 60: sinusoidal amplitude modulation detection at 60 Hz, FM 2 Hz: frequency modulation detection at 2 Hz, FM 10 Hz: frequency modulation detection at 10 Hz. For SAM 60 Hz, pitch discrimination (6000 Hz), FM 2 Hz and FM 10 Hz scores are reported as thresholds for discrimination and thus lower scores indicate better performance. For all other tests, scores are reflect number of correct responses, and thus higher scores are better.

Measure	Schizophrenia		Controls		Cohen's d
	Mean	SD	Mean	SD	
CATS	78.09**	11.35	83.65	10.85	0.50
MAV	41.18**	5.87	44.71	5.51	0.62
Emotion composite	47.51**	9.77	53.27	9.41	0.60
BACS composite	33.93**	15.73	43.78	14.45	0.65
UPSA total	36.07**	5.11	38.35	4.41	0.47
Frequency disc.	29.34**	30.83	46.00	32.25	0.53
Intensity disc.	25.10**	29.68	44.96	30.88	0.66
Duration disc.	23.93**	29.58	40.07	34.02	0.51
Pulse-Train	25.03**	29.64	41.53	34.57	0.51
Embedded Tone	19.55**	27.70	36.15	31.00	0.57
Temporal Order	21.49	27.76	28.25	29.48	0.24
Syllable Order	11.78**	16.39	23.86	26.63	0.56
Syllable Recognition	31.15*	27.47	43.04	29.92	0.41
Formant discrimination	47.72**	9.30	52.92	10.14	0.53
SAM60	3.01**	0.85	2.56	0.81	0.54
Pitch discrimination (6000 Hz)	5.53	1.87	5.04	1.78	0.27
FM 2 Hz	0.68*	0.29	0.57	0.23	0.41
FM 10 Hz	0.22	0.11	0.19	0.08	0.31

  * p < .05.
  ** p < .01.

3.4. Relationships between basic auditory processing and cognition, emotion recognition, functional capacity and clinical symptoms

All tests of basic auditory processing correlated significantly with all tests of cognition, with the exception of the Token Motor task, which exhibited the lowest correlations with basic auditory processing (Table 3).

All tests of basic auditory processing correlated with both batteries of emotion recognition (Table 4). In all instances, the correlations were stronger with the CATS than with the MAV. To remove the effects of cognition, we computed partial correlations, controlling for BACS Symbol Coding performance. While the BACS does produce a composite score, several of the tests that comprise the battery contain auditory components, and thus controlling for cognition using the composite score might lead to overcorrection and spuriously reduce the correlation between audition and emotion recognition. The BACS Symbol Coding test was chosen to control for cognition because it is a good proxy for

Table 4
Correlations between basic auditory skills and emotion recognition. CATS: Comprehensive Affect Testing System Prosody Scale, MAV: Montreal Affective Voices, emotion composite: composite score of CATS and MAV, SAM 60: sinusoidal amplitude modulation detection at 60 Hz, FM 2 Hz: frequency modulation detection at 2 Hz, FM 10 Hz: frequency modulation detection at 10 Hz.
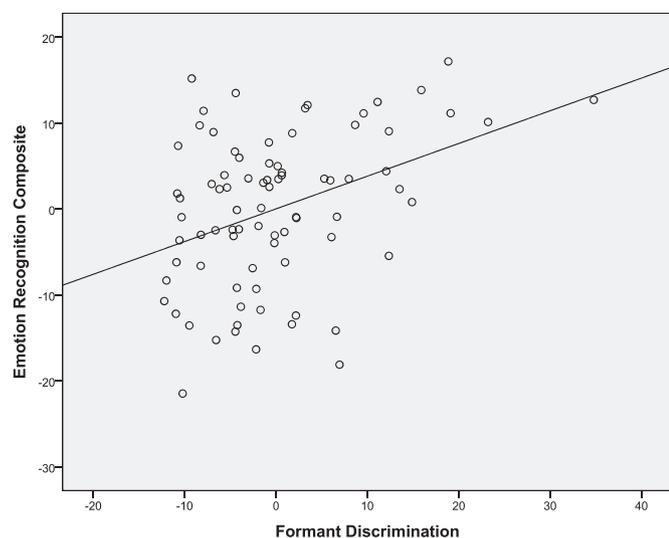
	CATS	MAV	Emotion comp	Emotion partial
Frequency	0.48**	0.35**	0.45**	0.31**
Intensity	0.46**	0.32**	0.42**	0.27*
Duration	0.42**	0.37**	0.43**	0.34**
Pulse-Train	0.37**	0.28**	0.35**	0.26*
Embedded Tone	0.41**	0.39**	0.45**	0.34**
Temporal Order	0.40**	0.33**	0.40**	0.29**
Syllable Order	0.31**	0.28*	0.30**	0.20
Syllable Recognition	0.46**	0.45**	0.49**	0.37**
Formant frequency	0.47**	0.42**	0.47**	0.38**
SAM60	−0.50**	−0.43**	−0.51**	−0.43**
Pitch 6000 Hz	−0.39**	−0.29*	−0.38**	−0.28*
FM 2 Hz	−0.42**	−0.30**	−0.40**	−0.35**
FM 10 Hz	−0.45**	−0.23*	−0.39**	−0.31**

  * p < .05.
  ** p < .01.

overall cognition and because it does not require audition and thus should not confound the effects of audition on emotion recognition (Dickinson et al., 2007). With the exception of Syllable Order, all correlations remained significant after controlling for cognition. Our a priori hypothesis that formant discrimination is associated with deficits in emotion recognition in patients was supported by this analysis as this association remained significant after controlling for cognition (Fig. 1).

To determine the ability of basic auditory skills performance to improve prediction of emotion recognition beyond what is accounted for by cognition, we performed a step-wise regression after controlling for cognition (by entering BACS Symbol Coding as the first step in the regression). Formant discrimination was the only basic auditory skill to enter the step-wise model of emotion discrimination after controlling for cognition. Formant discrimination accounted for an additional 14.5% of variance in emotion identification beyond that accounted for by cognition and rendered the model a significant predictor of emotion recognition ($F_{2,59} = 7.40$, p < .01), while the model including only cognition was not. To quantify the additional variance accounted for by formant discrimination over pitch processing (the most widely investigated association to date), we performed a multiple-regression, entering frequency discrimination at 1000 Hz in the first step and then entering formant discrimination in the second step. We found that formant discrimination accounted for an additional 6.9% variance in emotion recognition over frequency discrimination alone ($F_{2,83} = 9.48$, p < .01).

Table 3
Correlation between basic auditory skills and cognition. BACS SC: Brief Assessment of Cognition in Schizophrenia – Symbol Coding, SAM 60: sinusoidal amplitude modulation detection at 60 Hz, FM 2 Hz: frequency modulation detection at 2 Hz, FM 10 Hz: frequency modulation detection at 10 Hz.

	Verbal memory	Digit seq.	Token motor	Sem. fluency	BACS SC	Tower of London	Composite
Frequency	0.40**	0.48**	0.25*	0.41**	0.48**	0.43**	0.55**
Intensity	0.40**	0.40**	0.22*	0.38**	0.52**	0.32**	0.51**
Duration	0.46**	0.44**	0.21	0.42**	0.33**	0.30**	0.48**
Pulse-Train	0.47**	0.30**	0.13	0.30**	0.30**	0.33**	0.41**
Embedded Tone	0.37**	0.42**	0.05	0.33**	0.38**	0.34**	0.42**
Temporal Order	0.41**	0.43**	0.28*	0.48**	0.36**	0.46**	0.54**
Syllable Order	0.41**	0.45**	0.21	0.54**	0.30**	0.43**	0.51**
Syllable Recognition	0.60**	0.58**	0.30**	0.48**	0.45**	0.33**	0.60**
Formant frequency	0.43**	0.55**	0.15	0.46**	0.35**	0.36**	0.51**
SAM60	−0.32**	−0.49**	−0.14	−0.42**	−0.33**	−0.41**	−0.50**
Pitch 6000 Hz	−0.35**	−0.41**	−0.19	−0.38**	−0.34**	−0.31*	−0.46**
FM 2 Hz	−0.26*	−0.32**	−0.20	−0.24*	−21	−0.38**	−0.38**
FM 10 Hz	−0.25*	−0.33**	−0.20	−0.22*	−31**	−0.25*	−0.38**

  * p < .05.
  ** p < .01.

Fig. 1. Partial regression plot of formant discrimination after controlling for cognition.

All tests of basic auditory capabilities correlated with UPSA total score, with Pearson correlation coefficients ranging from 0.28 (FM 2 Hz) to 0.53 (Formant Discrimination) (Supplemental Table 1). Comparing the domains of functional capacity, it is clear that auditory skills correlated less strongly to the domain of comprehension/planning than other domains. Also, the auditory skills of pitch discrimination at 6000 Hz, pulse train discrimination and FM detection (at both 2 Hz and 10 Hz) appeared to correlate less strongly with functional capacity than the other basic auditory skills.

Correlations of basic auditory capabilities with symptoms were generally quite low (Supplemental Table 2). With only 2 exceptions, basic auditory capabilities correlated more strongly with negative symptoms than with positive or general symptoms.

4. Discussion

This study provides the first comprehensive assessment of basic auditory processing in individuals with schizophrenia and the relationships between these basic auditory skills and auditory emotion processing, cognition and symptoms. Patients demonstrated impaired ability to recognize emotion in voice with a medium effect size, similar to deficits reported in other studies (Gold et al., 2012; Kantrowitz et al., 2013; Leitman et al., 2005). Like other studies, we found that patients with schizophrenia are impaired in their ability to distinguish differences in lower frequency pitches and that this ability correlated with their ability to understand emotion in voice. Unlike other studies, we tested a wide variety of auditory abilities spanning all empirically defined domains of auditory function and included complex and vocally-based stimuli. We found that auditory impairments are widespread and involve all domains of auditory function, with mostly medium effect sizes. The magnitude of these impairments was very similar to the magnitude of the cognitive impairments in this group of patients compared to healthy controls. It should be noted that our healthy control group, which was recruited to rigorously match our patient sample on age, sex, race, and parental education, scored more poorly on the BACS cognitive battery than our previous normative study (Keefe et al., 2008), registering an age- and sex-corrected composite T-score of 43.78. Healthy controls were recruited through use of flyers throughout the Durham area and through Craigslist advertising. Subjects needed to be available for a full day of testing during normal business hours, which may have biased our sample towards lower cognitively functioning individuals.

Previous studies in patients with schizophrenia have consistently demonstrated a correlation between pitch processing of simple sine waves and emotion recognition from vocal tone (Holcomb et al., 1995; Javitt et al., 1997; Kantrowitz et al., 2015; Leitman et al., 2005; Leitman et al., 2010; March et al., 1999). Like previous studies, we found a relationship between pitch processing of lower frequency (1000 Hz) sine wave tones and emotion recognition. However, in our study we assessed a wide array of basic auditory abilities, allowing us to investigate whether this relationship between pitch processing and emotion recognition is unique or if a wider assortment of basic auditory impairments is associated with emotion recognition deficits in patients with schizophrenia. Indeed, correlational analyses in patients indicated widespread association between basic auditory impairment and emotion recognition deficits. We found the relationships of auditory skills with emotion recognition to be stronger with the CATS than with the MAV. We speculate that this may be due to the longer duration of the CATS stimuli and because the CATS stimuli are verbal as opposed to non-verbal affective bursts employed by the MAV. Impairment on all basic auditory skills, except syllable order recognition, was significantly correlated with emotion recognition deficits after removing the effects of cognitive impairment. These results suggest that the auditory perception impairment is far more prevalent than previously demonstrated.

While our results primarily indicate that the impairments in basic auditory skills are broad ranging, and that most of these impairments contribute to emotion recognition deficits, there is some evidence that particular auditory skills may be more central to patients' difficulty in recognizing emotion. Other studies have demonstrated pitch processing deficits in patients at base frequencies between 500 Hz and 2000 Hz, in the range (<4000 Hz) at which cells in auditory brain structures are able to phase-lock their response to periodic components of an auditory stimulus (Holcomb et al., 1995; Javitt et al., 1997; March et al., 1999; Rose et al., 1967). This phase-locked signal allows for extra precision of frequency encoding over a purely place-coded signal. To test if frequency encoding is impaired when phase-locking is not involved, we included a pitch processing test at a base frequency of 6000 Hz (outside of the phase-locking range). While our patients were impaired in their ability to distinguish pitches at 1000 Hz, frequency processing at 6000 Hz was one of only three skills that was found to be intact. Another of the intact skills, detection of frequency modulation at 10 Hz involved stimuli that elicit minimal phase-locking compared to the same task at 2 Hz, for which patients did demonstrate significant impairment. Similarly, the four strongest correlations between basic auditory processing and emotion recognition, after controlling for cognition, emphasize frequency processing in the phase-locking range (SAM60, Formant Discrimination, Syllable Recognition and FM 2 Hz). Also of note, Formant Discrimination and Syllable Recognition were the two tests in our battery that involved discrimination of the most naturalistic, speech-like sounds. The fact that they were among the most highly correlated with emotion recognition suggests that discrimination of complex, speech sounds may be a more important skill than discrimination of stimuli based on simple, artificial sounds. Formant discrimination deficits, specifically, accounted for unique variance in emotion recognition impairment beyond general cognitive impairment and basic pitch processing deficits in patients with schizophrenia, suggesting that the inability to perceive vowel sounds in schizophrenia patients contributes to their inability to understand others' emotions.

Impairments in emotion recognition are a key contributor to poor functional outcomes in schizophrenia (Brekke et al., 2005). Since the data from the current study suggest that widespread basic auditory processing deficits are associated with emotion recognition deficits in schizophrenia, a remediative approach that includes training on all basic auditory processing domains may be optimal for improving emotion recognition. Further, our results suggest that these auditory processing deficits may relate to deficits in functional capacity as measured by the UPSA. Thus, remediation that improves these basic auditory skills may have the potential to improve functional capacity as well. Tasks emphasizing spectral processing of naturalistic stimuli in the phase locking range might be especially helpful. These types of

natural vocal stimuli may be more relevant than artificial stimuli to the distinctions that one must make when judging a speaker's emotion. The formant discrimination task we employed in this study fulfills these criteria. Healthy individuals are remarkably adept at judging a speaker's emotion from very brief exposure to vowel segments of speech. Also, sensitivity to changes in vowel formant frequencies can be taught in healthy individuals (Kewley-Port, 2001), suggesting that the deficits exhibited by patients in this skill could be remediated.

4.1. Limitations

Due to the low performance of our control group, the cognitive effect size between our groups was smaller than has previously been reported. It is possible that higher effect sizes on the basic auditory and emotion recognition batteries would have been observed with a higher functioning control group. Also, because this study used a cross-sectional design, we were unable to assess the relationship between basic auditory deficits and emotion recognition deficits over time. Although we did not correct for multiple comparisons, the sheer number of findings supports the conclusion that schizophrenia is associated with widespread impairment in basic auditory processing and auditory emotion recognition. Finally, because patients were medicated, it is possible that medications could be impairing both basic auditory function and emotion recognition.

4.2. Conclusions

Patients with schizophrenia exhibit widespread deficits in basic auditory impairments associated with difficulty identifying emotions from the acoustic properties of a speaker's voice. While many basic auditory skills contribute to this impairment, deficient formant discrimination appears to be a key skill contributing to impaired emotion recognition as this was the only basic auditory skill to enter a step-wise multiple regression after first entering a measure of cognitive impairment, and formant discrimination accounted for significant unique variance in emotion recognition performance after accounting for deficits in pitch processing.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.schres.2018.08.031.

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.

Badcock, J.C., 2010. The cognitive neuropsychology of auditory hallucinations: a parallel auditory pathways framework. Schizophr. Bull. 36 (3), 576–584.

Banse, R., Scherer, K.R., 1996. Acoustic profiles in vocal emotion expression. J. Pers. Soc. Psychol. 70 (3), 614–636.

Bowie, C.R., Reichenberg, Abraham, Patterson, Thomas L., Heaton, Robert K., Harvey, Philip D., 2006. Determinants of real-world functional performance in schizophrenia subjects: correlations with cognition, functional capacity, and symptoms. Am. J. Psychiatry 163 (3), 418–425.

Brekke, J., Kay, D.D., Lee, K.S., Green, M.F., 2005. Biosocial pathways to functional outcome in schizophrenia. Schizophr. Res. 80 (2–3), 213–225.

Chandrasekaran, B., Hornickel, J., Skoe, E., Nicol, T., Kraus, N., 2009. Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: implications for developmental dyslexia. Neuron 64 (3), 311–319.

Cienfuegos, A., March, L., Shelley, A.-M., Javitt, D.C., 1999. Impaired categorical perception of synthetic speech sounds in schizophrenia. Biol. Psychiatry 45 (1), 82–88.

Corcoran, C.M., Keilp, J.G., Kayser, J., Klim, C., Butler, P.D., Bruder, G.E., Gur, R.C., Javitt, D.C., 2015. Emotion recognition deficits as predictors of transition in individuals at clinical high risk for schizophrenia: a neurodevelopmental perspective. Psychol. Med. 45 (14), 2959–2973.

Cornblatt, B.A., Carrión, R.E., Addington, J., Seidman, L., Walker, E.F., Cannon, T.D., Cadenhead, K.S., McGlashan, T.H., Perkins, D.O., Tsuang, M.T., Woods, S.W., Heinssen, R., Lencz, T., 2012. Risk Factors for Psychosis: Impaired Social and Role Functioning. Schizophr. Bull. 38 (6), 1247–1257.

Dickey, C.C., Morocz, I.A., Niznikiewicz, M.A., Voglmaier, M., Toner, S., Khan, U., Dreusicke, M., Yoo, S.-S., Shenton, M.E., McCarley, R.W., 2008. Auditory processing abnormalities in schizotypal personality disorder: an fMRI experiment using tones of deviant pitch and duration. Schizophr. Res. 103 (1–3), 26–39.

Dickey, C.C., Morocz, I.A., Minney, D., Niznikiewicz, M.A., Voglmaier, M.M., Panych, L.P., Khan, U., Zacks, R., Terry, D.P., Shenton, M.E., McCarley, R.W., 2010. Factors in sensory processing of prosody in schizotypal personality disorder: an fMRI experiment. Schizophr. Res. 121 (1–3), 75–89.

Dickinson, D., Ramsey, M.E., Gold, J.M., 2007. Overlooking the obvious: a meta-analytic comparison of digit symbol coding tasks and other cognitive measures in schizophrenia. Arch. Gen. Psychiatry 64 (5), 532–542.

Franklin, M.S., Sledge Moore, K., Yip, C.-Y., Jonides, J., Rattray, K., Moher, J., 2008. The effects of musical training on verbal memory. Psychol. Music 36 (3), 353–365.

Fukui, I., Burger, R.M., Ohmori, H., Rubel, E.W., 2010. GABAergic inhibition sharpens the frequency tuning and enhances phase locking in chicken nucleus magnocellularis neurons. J. Neurosci. 30 (36), 12075–12083.

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.

Green, M.F., Schooler, N.R., Kern, R.S., Frese, F.J., Granberry, W., Harvey, P.D., Karson, C.N., Peters, N., Stewart, M., Seidman, L.J., Sonnenberg, J., Stone, W.S., Walling, D., Stover, E., Marder, S.R., 2011. Evaluation of functionally meaningful measures for clinical trials of cognition enhancement in schizophrenia. Am. J. Psychiatry 168 (4), 400–407.

Grisham, J.R., Fullana, M.A., Mataix-Cols, D., Moffitt, T.E., Caspi, A., Poulton, R., 2011. Risk factors prospectively associated with adult obsessive–compulsive symptom dimensions and obsessive–compulsive disorder. Psychol. Med. 41 (12), 2495–2506.

Hoekert, M., Kahn, R., 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.

Javitt, D.C., 2009. When doors of perception close: bottom-up models of disrupted cognition in schizophrenia. Annu. Rev. Clin. Psychol. 5 (1), 249–275.

Javitt, D.C., Strous, R.D., Grochowski, S., Ritter, W., Cowan, N., 1997. Impaired precision, but normal retention, of auditory sensory ("echoic") memory information in schizophrenia. J. Abnorm. Psychol. 106 (2), 315–324.

Johnson, K.L., Nicol, T.G., Kraus, N., 2005. Brain stem response to speech: a biological marker of auditory processing. Ear Hear. 26 (5), 424.

Johnson, K.L., Nicol, T., Zecker, S.G., Bradlow, A.R., Skoe, E., Kraus, N., 2008. Brainstem encoding of voiced consonant–vowel stop syllables. Clin. Neurophysiol. 119 (11), 2623–2635.

Kantrowitz, J.T., Leitman, D.I., Lehrfeld, J.M., Laukka, P., Juslin, P.N., Butler, P.D., Silipo, G., Javitt, D.C., 2013. Reduction in tonal discriminations predicts receptive emotion processing deficits in schizophrenia and schizoaffective disorder. Schizophr. Bull. 39 (1), 86–93.

Kantrowitz, J.T., Hoptman, M.J., Leitman, D.I., Moreno-Ortega, M., Lehrfeld, J.M., Dias, E., Sehatpour, P., Laukka, P., Silipo, G., Javitt, D.C., 2015. Neural substrates of auditory emotion recognition deficits in schizophrenia. J. Neurosci. 35 (44), 14909–14921.

Kay, S.R., Fiszbein, A., Opler, L.A., 1987. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr. Bull. 13 (2), 261–276.

Keefe, R.S.E., Goldberg, T.E., Harvey, P.D., Gold, J.M., Poe, M.P., Coughenour, L., 2004. The Brief Assessment of Cognition in Schizophrenia: reliability, sensitivity, and comparison with a standard neurocognitive battery. Schizophr. Res. 68 (2–3), 283–297.

Keefe, R.S.E., Harvey, P.D., Goldberg, T.E., Gold, J.M., Walker, T.M., Kennel, C., Hawkins, K., 2008. Norms and standardization of the Brief Assessment of Cognition in Schizophrenia (BACS). Schizophr. Res. 102 (1–3), 108–115.

Kewley-Port, D., 2001. Vowel formant discrimination II: effects of stimulus uncertainty, consonantal context, and training. J. Acoust. Soc. Am. 110 (4), 2141–2155.

Kidd, G.R., Watson, C.S., Gygi, B., 2007. Individual differences in auditory abilities. J. Acoust. Soc. Am. 122 (1), 418–435.

Kraus, N., Chandrasekaran, B., 2010. Music training for the development of auditory skills. Nat. Rev. Neurosci. 11 (8), 599–605.

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., Hoptman, M.J., Foxe, J.J., Saccente, E., Wylie, G.R., Nierenberg, J., Jalbrzikowski, M., Lim, K.O., Javitt, D.C., 2007. The neural substrates of impaired prosodic detection in schizophrenia and its sensorial antecedents. Am. J. Psychiatry 164 (3), 474–482.

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.

Liu, C., Kewley-Port, D., 2004. Vowel formant discrimination for high-fidelity speech. J. Acoust. Soc. Am. 116 (2), 1224–1233.

March, L., Cienfuegos, A., Goldbloom, L., Ritter, W., Cowan, N., Javitt, D.C., 1999. Normal time course of auditory recognition in schizophrenia, despite impaired precision of the auditory sensory ("echoic") memory code. J. Abnorm. Psychol. 108 (1), 69–75.

Niemi, L.T., Suvisaari, J.M., Haukka, J.K., Lönnqvist, J.K., 2005. Childhood predictors of future psychiatric morbidity in offspring of mothers with psychotic disorder. Br. J. Psychiatry 186 (2), 108–114.

Nikjeh, D.A., Lister, J.J., Frisch, S.A., 2009. Preattentive cortical-evoked responses to pure tones, harmonic tones, and speech: influence of music training. Ear Hear. 30 (4), 432–446. https://doi.org/10.1097/AUD.1090b1013e3181a1061bf1092.

Rose, J.E., Brugge, J.F., Anderson, D.J., Hind, J.E., 1967. Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. J. Neurophysiol. 30 (4), 769–793.

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

Sachs, G., Winklbaur, B., Jagsch, R., Lasser, I., Kryspin-Exner, I., Frommann, N., Wölwer, W., 2012. Training of affect recognition (TAR) in schizophrenia—impact on functional outcome. Schizophr. Res. 138 (2), 262–267.

Schaffer, S.G., Froming, K.B., Gregory, A.L., Levy, C.M., Ekman, P., 2006. Emotion Processing: The Comprehensive Affective Testing System User's Manual. Psychology Software Inc.

Sek, A., Moore, B.C.J., 1995. Frequency discrimination as a function of frequency, measured in several ways. J. Acoust. Soc. Am. 97 (4), 2479–2486.

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 (1–3), 214–220.

Teale, P., Collins, D., Maharajh, K., Rojas, D.C., Kronberg, E., Reite, M., 2008. Cortical source estimates of gamma band amplitude and phase are different in schizophrenia. NeuroImage 42 (4), 1481–1489.

Thompson, A.D., Bartholomeusz, C., Yung, A.R., 2011. Social cognition deficits and the 'ultra high risk' for psychosis population: a review of literature. Early Interv. Psychiatry 5 (3), 192–202.

Waaramaa, T., Laukkanen, A.-M., Airas, M., Alku, P., 2010. Perception of emotional valences and activity levels from vowel segments of continuous speech. J. Voice 24 (1), 30–38.