Psychiatry Research 205 (2013) 192–198

Contents lists available at SciVerse ScienceDirect

Psychiatry Research

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

Prosody recognition and audiovisual emotion matching in schizophrenia: The contribution of cognition and psychopathology

Filomena Castagna [a], Cristiana Montemagni [a], Anna Maria Milani [a], Giuseppe Rocca [a], Paola Rocca [a,*], Massimo Casacchia [b], Filippo Bogetto [a]

[a] Department of Neuroscience, Psychiatric Section, University of Turin, Via Cherasco, 11, 10126 Turin, Italy
[b] Department of Science of Health, University of L'Aquila, Via Vetoio Coppito 2, 67010 L'Aquila, Italy

ARTICLE INFO

Article history:
Received 14 September 2011
Received in revised form
16 August 2012
Accepted 25 August 2012

Keywords:
Schizophrenia
Prosody
Audiovisual emotion matching
Cognition
Psychopathology

ABSTRACT

This study aimed to evaluate the ability to decode emotion in the auditory and audiovisual modality in a group of patients with schizophrenia, and to explore the role of cognition and psychopathology in affecting these emotion recognition abilities. Ninety-four outpatients in a stable phase and 51 healthy subjects were recruited. Patients were assessed through a psychiatric evaluation and a wide neuropsychological battery. All subjects completed the Comprehensive Affect Testing System (CATS), a group of computerized tests designed to evaluate emotion perception abilities. With respect to the controls, patients were not impaired in the CATS tasks involving discrimination of nonemotional prosody, naming of emotional stimuli expressed by voice and judging the emotional content of a sentence, whereas they showed a specific impairment in decoding emotion in a conflicting auditory condition and in the multichannel modality. Prosody impairment was affected by executive functions, attention and negative symptoms, while deficit in multisensory emotion recognition was affected by executive functions and negative symptoms. These emotion recognition deficits, rather than being associated purely with emotion perception disturbances in schizophrenia, are affected by core symptoms of the illness.

1. Introduction

Social cognition abilities are compromised in individuals with schizophrenia (Green et al., 2008) and these impairments are linked to a deficient level of social functioning (Couture et al., 2006; Fett et al., 2010). One of the most impaired functions is the ability to decode emotions based on either facial expression and speech intonation. While a considerable number of studies have consistently shown evidence that the recognition of emotional faces is disturbed in patients with schizophrenia (Trémeau, 2006), studies on prosody discrimination and multisensory integration of facial and vocal affect (audiovisual emotion matching) are rare.

The term "prosody" discrimination refers to the ability to recognize, comprehend, and produce affect as well as semantic meaning based on the intonation, stress, and rhythm patterns of vocal utterances. Emotional prosody recognition refers to the ability to detect affect and to infer emotion based on prosodic information, while nonemotional prosody recognition refers to the ability to extract meaning from the intonational contour, for example, in differentiating interrogative sentences from declarative sentences. Nonemotional prosody seems to be preserved in patients with schizophrenia (Murphy and Cutting, 1990; Pijnenborg et al., 2007), while a relatively stable deficit in emotional prosody has been demonstrated (Hoekert et al., 2007).

Studies examining the associations between impaired recognition of emotional prosody and symptoms have let to inconclusive results. While some studies report a relationship between impaired emotional prosody and positive symptoms, other studies report either a relationship with negative symptoms or no relationship at all (for reviews see Bozikas et al., 2004; Trémeau, 2006).

It has been argued that cognitive deficit in schizophrenia could underlie an impaired perception of emotional prosody, but results with regard to the covariation of general cognitive functioning and prosody perception are not consistent. Emotional prosody perception was correlated with neurocognitive measures, particularly with performance on tests of attention and executive functions (Bozikas et al., 2004; Pijnenborg et al., 2007; Scholten et al., 2008; Dickey et al., 2010); however, this association has not been found in all studies (Whittaker et al., 1994). In addition, basic pitch perception deficits may be associated with a deficit of emotional prosody (Leitman et al., 2005).

In spite of the importance of integrating information from different sensory modalities during social interactions, the field of multisensory integration in schizophrenia has not been

* Corresponding author. Tel.: +39 011 6336780; fax: +39 011 673473.
E-mail address: paola.rocca@unito.it (P. Rocca).

0165-1781/$ - see front matter © 2012 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.psychres.2012.08.038

systematically explored. Relatively few authors (de Gelder et al., 2005; de Jong et al., 2009, 2010) have examined how visual and auditory information are integrated, whether this integration process is impaired in schizophrenia and whether potential deficits lead to less adaptive behaviors. These studies have shown that patients with schizophrenia were significantly less able to integrate vocal and facial emotional information compared to healthy people. Moreover, using cross-modal emotional recognition tasks (emotion-congruent and -incongruent conditions), some authors have also found a cross-modal influence between faces and voices that can facilitate or reduce behavioral performance (de Gelder et al., 1995, 2005; de Jong et al., 2009, 2010). No study has previously investigated the relationship between audiovisual emotion matching, symptoms and cognitive functions.

The present study aimed to evaluate the ability to decode emotion in the auditory modality and in the audiovisual modality in a group of patients with stable schizophrenia. Specifically, aims and hypotheses of the present study, guided by previously reported literature, were as follows:

1. To explore whether individuals with stable schizophrenia are impaired in nonemotional and emotional prosody recognition tasks with respect to healthy individuals. The role of other factors related to schizophrenia and previously investigated as putative contributing factors to prosody processing in schizophrenia, such as cognition and psychopathology, has been analysed. Given the mixed results of previous studies, we intended this as an exploratory analysis and expected that cognition and psychopathology would significantly predict performances of patients in nonemotional and emotional prosody recognition tasks.
2. To explore whether individuals with stable schizophrenia are impaired in audiovisual emotion matching with respect to healthy individuals. As with aim 1, we intended to explore the effects of cognition and psychopathology in affecting patients' performances in audiovisual emotion matching tasks.

2. Methods

2.1. Participants

Patients were referred to the Psychiatric Section, Department of Neuroscience, University of Turin, and the Mental Health Department 1 South of Turin in the period between January 2007 and January 2008. Patients were initially evaluated by a clinician–psychiatrist, and if they met DSM-IV-TR (American Psychiatric Association, 2000) criteria for schizophrenia, they were evaluated subsequently by our research team. The diagnosis of schizophrenia was confirmed by two expert clinicians (P.R., C.M.), using the Structured Clinical Interview for DSM-IV Disorders (SCID) (First et al., 1997). Subjects were excluded if they had a current disorder other than schizophrenia on Axis I of the DSM-IV-TR, a current or past codiagnosis of autistic disorder or another pervasive developmental disorder, a history of severe head injury (coma $\geq$48 h) and a diagnosis of mental disorder due to a general medical condition.

At the time of study entry, patients had been clinically stable for at least 6 months as judged by the treating psychiatrist, i.e. during this period all patients had to be treated as outpatients, treatment regimen had not been modified, and there was no essential change in psychopathology. In addition to medical records, patients were considered to be in a stable phase as assessed from reports from the patients themselves, and observations of the psychiatric personnel, relatives and personnel in the psychiatric community.

Ninety-four consecutive outpatients who met the inclusion criteria were enrolled in the study. Patients were evaluated using a semistructured interview to assess demographic features. Data were collected to determine age, gender, education, age at onset of schizophrenia (report of first contact with a psychiatric service) and antipsychotic treatment. For each patient the dose equivalent to 100 mg/day of chlorpromazine was calculated (Woods, 2003).

Controls (N=51) were healthy individuals, recruited among clinical staff and medical students in the Department of Neuroscience, Psychiatric Section, University of Turin. The inclusion criteria for healthy controls included no current or past psychiatric diagnosis or treatment.

The study was carried out in accordance with the Declaration of Helsinki 1995 (as revised in Edinburgh 2000) and Good Clinical Practice. The protocol was approved by a Local Research Ethics Committee and written informed consent was obtained from all subjects after a complete description of the study.

2.2. Assessment instruments

2.2.1. Psychiatric assessment

In the group of patients, current severity of psychopathological symptoms was assessed using the Positive and Negative Syndrome Scale (PANSS) (Kay et al., 1987), which includes Positive Symptom, Negative Symptom, and General Psychopathology subscales.

In an attempt to reduce inter-rater variability, all raters were trained to administer the psychometric tools according to common standards. Also prior to the commencement of the present study, they participated in a pilot study in order to reach a consensus on ratings that were obtained using psychometric scales. The inter-rater reliability was 0.87–0.91 for the PANSS Positive, 0.83–0.87 for the PANSS Negative, and 0.88–0.92 for the PANSS General Psychopathology sub-scales. Efforts were made to maintain inter-rater reliability across the entire study period, including careful calibration and standardization procedures and regular, in-depth review of a sample of interviews with the lead author.

2.2.2. Neuropsychological assessment

Patients were assessed with a neuropsychological test battery by one trained psychologist (F.C.), who was unaware of clinical characteristics and results of psychiatric rating scales. The battery was administered and scored according to the established procedures for each test the day after the psychiatric assessments.

The assessment was fixed in one session of 1 h per patient. No subject was familiarized with the tests. The Stroop Test (Stroop, 1935) and the Trail MakingTest (TMT) (Reitan, 1958) were used to estimate attentive functions. The California Verbal Learning Test (CVLT; Delis et al., 1987) was used to assess verbal memory-learning. The Wisconsin Card Sorting Test (WCST) (Heaton et al., 1993) was used to assess abstraction and cognitive flexibility. Premorbid intelligence was been evaluated with Test di Intelligenza Breve (TIB) (Sartori et al., 1997), an equivalent version of the National Adult Reading Test (NART) (Nelson, 1982) for Italian people.

2.2.3. Prosody and audiovisual emotion recognition assessment

In order to assess the ability to recognize emotions expressed from facial expressions and human voices, the Comprehensive Affect Testing System (CATS) (Froming et al., ©2000–2006) was administered to patients and controls. The CATS is a new computerized assessment battery designed to evaluate emotion perception abilities. The subtests are designed to test facial and prosodic discrimination and identification, facial affect matching with and without verbal meaning, and prosodic processing with and without verbal denotation (and with conflicting or congruent semantic content). For the purpose of our analysis we chose to use only the tasks requiring the processing of emotional prosody and semantics. A male actor speaks on the prosodic subtests. Emotional stimuli included the following emotions: happiness, sadness, anger, surprise, disgust, fear, or neutral mood.

The methods involved in the development of the test, including reliability data, are discussed in the CATS manual (Schaffer et al., 2006), which is available on the publisher's website: http://www.psychologysoftware.com/CATS.htm. The instructions for the tasks were translated by our group in Italian language and group into Italian and used for the present research (the Italian version of the CATS can be requested at the website: http://www.psychologysoftware.com/testing instruments.htm).

The testing session was about 40 min and was conducted in a quiet room free from auditory and visual distractors. One trained psychologist (F.C.) administered these tasks, pressing the button on the computer after the subject's verbal answer, without time limits. The system collects the accuracy of each answer submitted (percentage of correct answers). A complete description of each selected subtest follows.

2.2.3.1. Nonemotional prosody recognition task

- *Nonemotional prosody discrimination* (six items): No faces are shown. A pair of non-affective sentences (e.g., the boy opened the window) is heard on each trial. In each trial, the same sentence is said twice, as a simple declarative sentence or as a question, or one of each in a random order (three possible pairs of sentences: declarative–declarative, declarative–question/question–declarative, question–question). The subject must indicate whether the sentences are said with the same or different intonation.

2.2.3.2. Emotional prosody recognition tasks

- *Name emotional prosody* (12 items): No faces are shown. One sentence is read at a time by the actor. The subject selects which emotion (happiness, sadness, anger, fright or neutrality) his voice expresses.
- *Conflicting prosody/meaning—attend to prosody* (12 items): No faces are shown. The subject is instructed to ignore the affective meaning represented in the