Abnormal Habituation of the Auditory Event-Related Potential P2 Component in Patients With Schizophrenia

Prune Mazer[1,2*†], Inês Macedo[1†], Tiago O. Paiva[1], Fernando Ferreira-Santos[1], Rita Pasion[1], Fernando Barbosa[1], Pedro Almeida[3], Celeste Silveira[4,5], Cassilda Cunha-Reis[1] and João Marques-Teixeira[1]

[1] Laboratory of Neuropsychophysiology, Faculty of Psychology and Education Sciences of the University of Porto, Porto, Portugal, [2] School of Health, Polytechnic Institute of Porto, Porto, Portugal, [3] Faculty of Law, School of Criminology and Interdisciplinary Research Center on Crime, Justice and Security, University of Porto, Porto, Portugal, [4] Faculty of Medicine, University of Porto, Porto, Portugal, [5] Psychiatry Department, Hospital S. João, Porto, Portugal

Auditory event-related potentials (ERP) may serve as diagnostic tools for schizophrenia and inform on the susceptibility for this condition. Particularly, the examination of N1 and P2 components of the auditory ERP may shed light on the impairments of information processing streams in schizophrenia. However, the habituation properties (i.e., decreasing amplitude with the repeated presentation of an auditory stimulus) of these components remain poorly studied compared to other auditory ERPs. Therefore, the current study used a roving paradigm to assess the modulation and habituation of N1 and P2 to simple (pure tones) and complex sounds (human voices and bird songs) in 26 first-episode patients with schizophrenia and 27 healthy participants. To explore the habituation properties of these ERPs, we measured the decrease in amplitude over a train of seven repetitions of the same stimulus (either bird songs or human voices). We observed that, for human voices, N1 and P2 amplitudes decreased linearly from stimulus 1–7, in both groups. Regarding bird songs, only the P2 component showed a decreased amplitude with stimulus presentation, exclusively in the control group. This suggests that patients did not show a fading of neural responses to repeated bird songs, reflecting abnormal habituation to this stimulus. This could reflect the inability to inhibit irrelevant or redundant information at later stages of auditory processing. In turn schizophrenia patients appear to have a preserved auditory processing of human voices.

Keywords: schizophrenia, auditory, EEG, event-related potentials, habituation, N1, P2

INTRODUCTION

The orienting response to novel/rare stimuli is considered a fundamental reaction in living organisms. However, when confronted with the repetition of stimuli in stable and non-threatening conditions, neural responses are expected to fade with repetition, demonstrating that humans tend to adapt to the surrounding environment (1). Auditory event-related potentials (ERP) and early components such as the N1 and P2, allow tracking the time course of neural activity related to auditory processing (2) and respective habituation

(i.e., component amplitudes decrease with successive repetitions) and sensory gating processes (i.e., inhibition of the processing of repeated information).

N1 is a negative component that peaks around 100 ms after the onset of an auditory stimulus and is followed by the P2, a positive component occurring approximately after 200 ms [3]. N1 and P2 are sensitive to the exogenous physical properties of the stimuli [4], although they also encompass endogenous characteristics [2, 3]. At the exogenous level, early auditory categorical differences have been reported for distinct categories of sounds (e.g., voices and bird songs) such that results suggest a rapid brain discrimination of human voices [5–8]. At the endogenous level, N1 also seems to reflect a selective-orienting attention response toward novel/rare stimuli [9, 10], while P2 probably represents a subsequent stage of stimulus recognition toward specific stimuli (e.g., human voices) independently of attentional demands [5–7, 11].

Regarding habituation processes, the N1-like orienting neural responses are likely to decrease in magnitude with the repeated presentation of an auditory stimulus [1, 9]. Similarly, P2 also shows habituation and sensory gating properties, especially for the consecutive stimuli repetition with shorter and constant inter-stimulus intervals [1, 4, 12]. Nevertheless, the mechanisms underlying ERP adaptation are complex [13] and further investigation is still required in order to unveil their potential clinical applications.

Several studies have focused indeed on the potential of auditory ERP components and habituation properties (e.g., P50 and P300) as diagnostic tools for schizophrenia [14, 15]. N1 and P2 can be particularly useful to study the mechanisms underlying the pathophysiology of schizophrenia since they are especially sensitive to early cortical auditory processing (that appears to be affected by this disorder) and seem to arise from temporal lobe generators, which are also affected in schizophrenia [9, 16]. More specifically, the ability to filter out irrelevant or repetitive information—as assessed through sensory gating and habituation processes—maps early pre-attentive and later-stages of information processing, respectively [17]. Research consistently reports reduced N1 amplitudes in first-episode, as well as chronic schizophrenic patients and first-degree relatives using both simple [18–21] and complex stimuli [22, 23]. Conversely, P2 is often examined in the context of the N1/P2 complex, and thus it has been less studied independently. Studies that have examined both components show either smaller P2 amplitudes [19] or no differences in first-episode schizophrenia (using simple stimuli) compared to healthy controls [21], despite systematic reductions in N1. A meta-analysis by Ferreira-Santos et al. [24] reviewed 20 studies on P2 in schizophrenia and showed that patients exhibited smaller amplitudes to standard stimuli and larger amplitudes for target stimuli when compared to controls in oddball tasks. In addition, even if fewer studies have assessed habituation processes for N1 and P2, there is evidence of less attenuation in schizophrenia [25, 26], although a meta-analysis by Rosburg [27] found that patients with schizophrenia display reduced N1 gating due to the initial/novel stimulus—not to effects of repetition/habituation. Such abnormalities extend to the processing of voice stimuli with N1 being smaller to

pre-recorded voices [23, 28] and P2 exhibiting larger amplitudes for voices with affective content [29] and no differences in amplitude for neutral content [30] in Schizophrenia.

Studying voice processing in schizophrenia is particularly important since auditory verbal hallucinations and abnormal social cognition represent well-recognized symptoms of this pathology. Also, the study of pre-attentive and later-stages of information processing could help us understand which stages of the auditory cortical information processing are impaired (and which are intact) in schizophrenia, as well as if the type of stimuli (i.e., complex, simple, or social) interferes with this process. Since neurobiological impairments precede the beginning of a full clinical syndrome, it is of extreme relevance to study the N1 and P2 in pre-psychotic and early-psychotic states of schizophrenia [31] to explore deficits underlying genetic susceptibility for schizophrenia [9, 20]. For this purpose, the current study assessed the habituation of auditory ERP components in first-episode schizophrenia such that amplitude attenuations at N1 and P2 time-windows were measured in a roving paradigm including both complex sounds (human voices and bird songs) and a target/rare pure sinusoidal tone. In healthy participants, the reviewed literature indicates that (a) N1 increases for rare stimuli vs. frequent stimuli [9, 10], (b) P2 is particularly responsive to human voices vs. bird sounds [5], and (c) both components decrease as a function of repetition/habituation [1, 4, 12, 32]. Inversely, we hypothesized that patients with schizophrenia would show: (H1) a diminished N1 amplitude for target stimuli—pure tones [9]; (H2) a diminished P2 to human voices, considering P2 specificity to voices [5] and deficits in voice processing in schizophrenia [33]; and (H3) reduced habituation of N1 and P2 when compared to controls [25], considering deficits reported in sensory gating in early ERP components in schizophrenia [34].

METHOD

Participants

Twenty-seven first-episode patients diagnosed with schizophrenia were originally recruited, but one was excluded due to excessive noise in EEG recordings. Thus, the clinical group included 26 participants (8 female; 2 left-handed), with a mean age of 27.88 years ($SD = 11.94$) and a mean education of 11.73 years ($SD = 2.69$). All patients were medicated with second generation antipsychotics corresponding to, on average, 12.63 mg of olanzapine equivalents ($SD = 5.01$). Mean illness duration was 15.11 months ($SD = 8.01$). The control group included 27 healthy individuals (seven female; two left-handed), with a mean age of 27.56 years ($SD = 7.79$) and a mean education of 13.67 years ($SD = 2.26$). All participants reported being free from head injuries, other neurological disorders, learning disabilities, or substance abuse, and all disclosed having normal hearing. There were statistically significant differences between groups for education ($p = 0.006$) but not for age ($p = 0.897$). Although education varied between groups, all participants in the patient group had completed at least compulsory school (9 years), except for one who completed 8 years. All participants signed the informed consent before the beginning of the experiment.

Stimuli and Procedures

The selection of auditory stimulus for the task was based on a preliminary validation study (for more details see section Validation study of the auditory stimuli in the **Supplementary Material**). Using a modified version of the roving standard frequency paradigm (35), auditory stimuli (i.e., bird songs, voices, and the pure tone) were delivered via headphones and E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA) with an intensity of 90 dB SPL.

The protocol included 16 complex auditory stimuli (i.e., eight bird songs and eight non-speech human voices) with a duration of 200 ms each (10 ms rise and fall times). The complex sounds were presented in fixed trains composed of seven repetitions of the same sound with an inter-stimulus interval of 1,000 ms. Each train was randomly repeated 5 times during the protocol, resulting in a total of 80 blocks. The target tone was randomly presented 40 times during the entire protocol (i.e., probability of ∼7%) and only appeared between trains (**Figure 1**). This target consisted of a 1,000 Hz pure tone with 70 ms duration (10 ms rise and fall times). During the task, a black screen with a fixation point was presented. Behavioral data was collected on the target tone with the main purpose of maintaining participants' attention and avoiding sleepiness and fatigue (which could impair EEG data quality due to alpha activity). Thus, participants were asked to pay attention and rapidly identify when the target tone was presented by pressing a button (i.e., in a SRBOX display; E-Prime).

EEG Data Collection and Pre-processing

EEG data were recorded with a digitizing rate of 500 Hz from 128 channels (Hydrocel Geodesic Sensor Net, NetAmps 300 amplifier, NetStation 4.2. software—Electrical Geodesics Inc.). The electrodes were referenced to the vertex (Cz) during recording and re-referenced offline to the average reference. Impedances were kept below 50 kΩ for all electrodes (high input impedance system). Data was pre-processed using EEGLAB (v11.4.2.2b) (36, 37) and custom MATLAB scripts. Continuous EEG records were downsampled to 250 Hz and band-pass filtered (0.2–30 Hz) and submitted to an Independent Components Analysis (ICA) decomposition after interpolation of bad channels (maximum of 10% per record). Correction of eye blink artifacts was carried out semi-automatically using the CORRMAP EEGLAB plug-in (38). We used a correlation threshold of 0.80 to identify the artifactual ICs and then subtracted their activity from the EEG data. Pre- and post-subtraction EEG traces were visually inspected to ensure that signals were not altered outside the time windows of eye blinks. EEG records were segmented into 1,000 ms epochs (−200 to 800 ms in peri-stimulus time). Epochs with voltages below −100 or above 100 μV were automatically rejected and all segments were subjected to visual inspection and manual artifact rejection. Finally, all epochs were baseline corrected (200 ms pre-stimulus) and averaged by condition for each participant.

ERP Analysis

Grand-averaged ERPs were computed and visually inspected to ensure that the expected ERP morphology was present. We analyzed frontal scalp regions in the left (i.e., FC5 cluster: E24, E27, E28, E29, E34, and E35) and right hemispheres (i.e., FC6 cluster: E110, E111, E116, E123, E124, and E117), where the frontotemporal positivity to voices has been shown to be maximal (5). Time-windows for the N1 and P2 components were based on the visual inspection of the grand-averages and on the typical settings described in the literature on auditory ERPs. Therefore, peak amplitudes were quantified as the minimum voltage in the 80–200 ms post-stimulus presentation time-window for N1 and as the maximum voltage in the 200–300 ms time-window for P2. Peaks were extracted for each participant at each channel and then averaged per train position (i.e., 1–7, for bird songs and voices) and stimulus type (i.e., bird songs, vocalization, and pure tone).

Statistical Analysis

For the analysis of behavioral data on the target detection task, responses were divided according to the sound (i.e., bird songs or voices) that was presented previous to the target (i.e., pure tone) and submitted to paired-sample t-tests to assess possible differences in reaction times (RTs). Additionally, independent samples t-tests were performed to assess group differences (i.e., patients vs. controls) in RTs (ms) and accuracy (percentage of times the target tone was correctly identified).

For neural data, separate Mixed Repeated Measure ANOVAs were conducted for N1 and P2. Pure tones (H1) were analyzed by hemisphere (within-participants factor: left, right) and group (between-participants factor: controls, patients). To replicate the conditions from previous studies that did not analyze the seven positions in the train (5), we explored group differences in the processing of complex stimulus (H2) only for the first stimulus of the train (within-participants factors: hemisphere—left, right, and stimulus category—bird songs, human voices). Finally, to explore the habituation properties of these ERPs for bird and human voices stimuli, we measured the decrease in amplitude across stimulus repetition in each train from S1 to S7 using linear contrasts (H3). As such this model entered with between groups and within-participants factors (train sequence—stimuli repetition from 1 to 7; stimulus category—voices, bird songs; hemisphere—left, right). All the pairwise comparisons were corrected for multiple comparisons using the Bonferroni Correction, and effect sizes were calculated (i.e., partial eta squared and Cohen's d).

RESULTS

Behavioral Results: Target/ Pure-Tone Detection

RTs were significantly longer for targets following the presentation of human voices ($M = 423.70$, $SD = 93.78$) compared to targets following bird songs ($M = 392.60$, $SD = 80.15$), $t_{(50)} = -5.26$, $p < 0.001$, $d = 0.737$. For targets following the presentation of bird songs, patients had significantly longer RTs ($M = 419.14$, $SD = 80.85$) compared to controls ($M = 364.98$, $SD = 70.2$), $t_{(49)} = 2.54$, $p = 0.014$, $d = 0.713$. Patients also had significantly longer RTs for targets following human
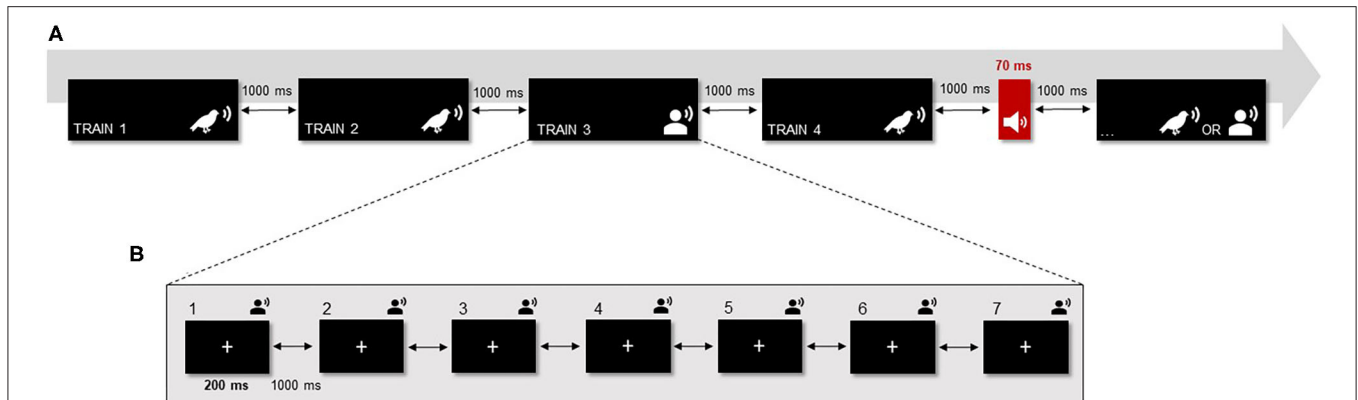
FIGURE 1 | Modified version of the roving standard frequency paradigm (Baldeweg et al., 2004). (A) There were 16 complex auditory stimuli divided into two categories: eight bird songs and eight human voices (non-speech voices). Each complex sound was presented in trains composed of seven repetitions of the same sound (ISI = 1,000 ms). Each train was randomly repeated five times during the protocol, resulting in a total of 80 blocks. Additionally, a target tone (i.e., a pure tone with duration of 70 ms) was randomly presented 40 times in between the trains of complex sounds. (B) Within the trains, each sound was presented for 200 ms (ISI = 1,000 ms) while a black screen with a fixation point was displayed.
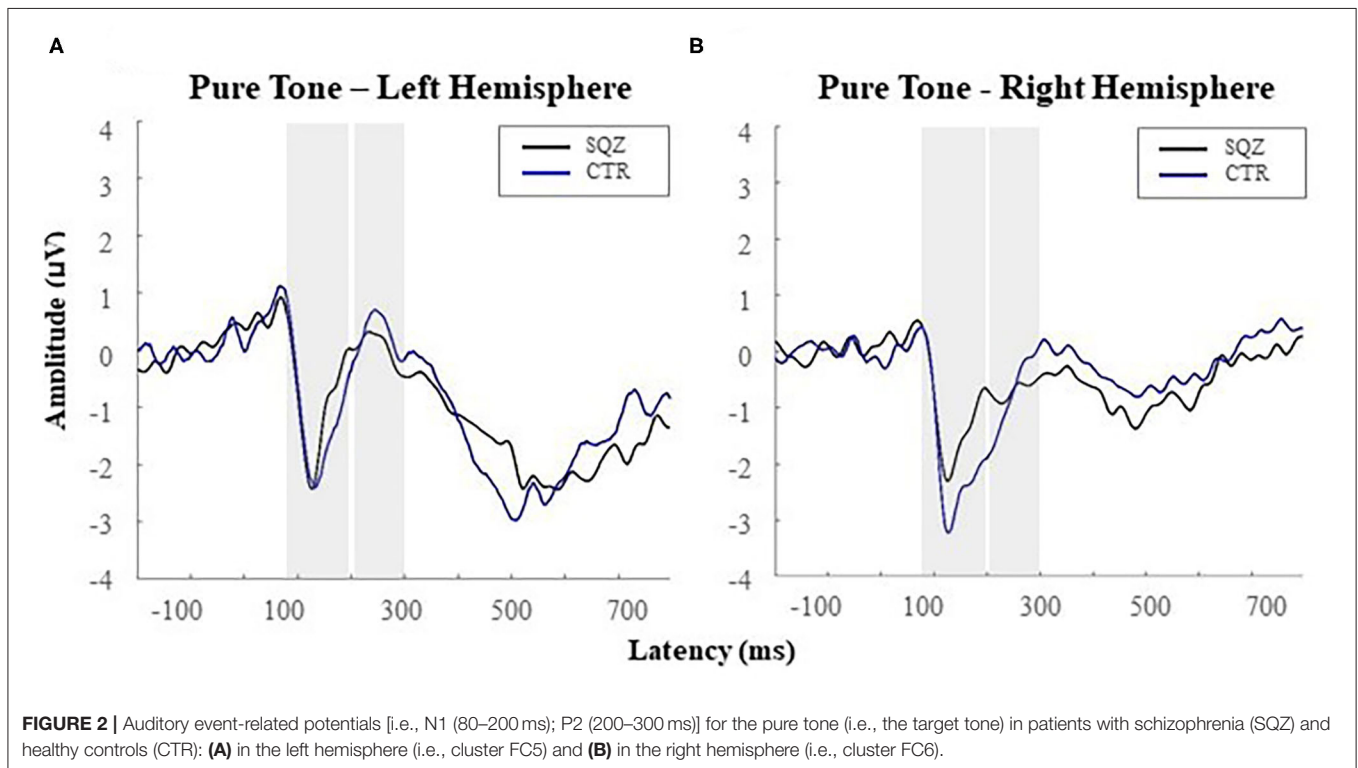

FIGURE 2 | Auditory event-related potentials [i.e., N1 (80–200 ms); P2 (200–300 ms)] for the pure tone (i.e., the target tone) in patients with schizophrenia (SQZ) and healthy controls (CTR): (A) in the left hemisphere (i.e., cluster FC5) and (B) in the right hemisphere (i.e., cluster FC6).

voices ($M = 461.95$, $SD = 92.71$) compared to controls ($M = 383.92$, $SD = 78.32$), $t_{(49)} = 3.24$, $p = 0.002$, $d = 0.909$.

The differences in accuracy between patients ($M = 96.50\%$, $SD = 7.01$) and controls ($M = 99.19\%$, $SD = 2.19$) were not statistically different, $t_{(51)} = -1.87$, $p = 0.067$, $d = 0.518$.
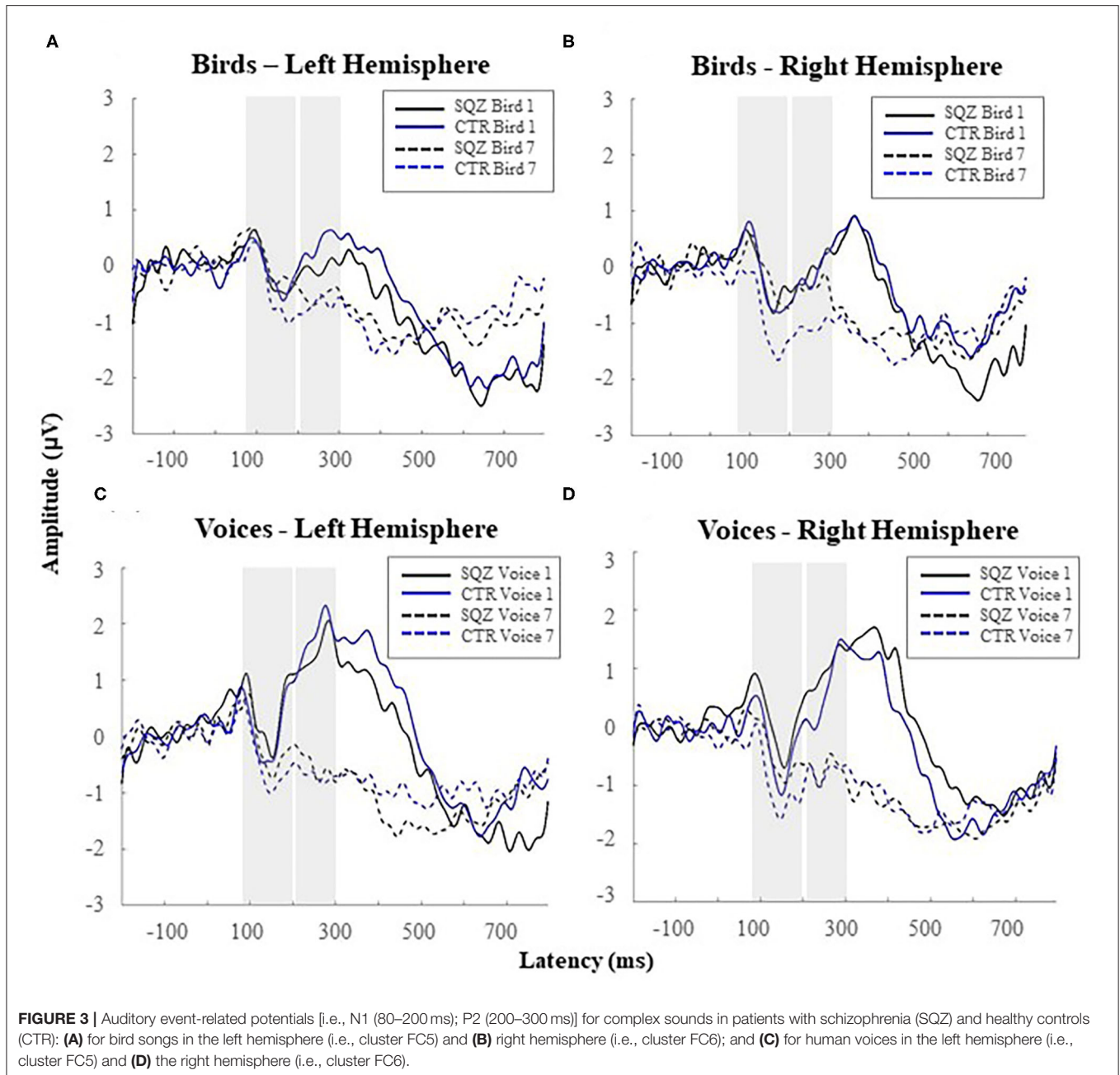
H1: Category— Simple Sound (Pure Tones—Target Stimuli)

At the N1 time-window, a significant hemisphere*group interaction was found, $F_{(1,51)} = 4.52$, $p = 0.038$, $\eta^2_p = 0.081$, revealing lower N1 amplitudes in patients ($M = -3.02$,

$SD = 1.48$) compared to controls ($M = -4.42$, $SD = 1.87$) in the right hemisphere, $t_{(51)} = -3.01$, $p = 0.024$, $d = 0.83$ (see **Figure 2**). No other significant effects were found for the N1.

For P2, a significant effect was found for hemisphere, $F_{(1,51)} = 5.35$, $p = 0.025$, $\eta^2_p = 0.095$, indicating higher amplitudes for the left ($M = 1.82$, $SD = 1.89$) comparing to the right hemisphere ($M = 1.02$, $SD = 2.12$), $t_{(52)} = 2.34$, $p = 0.023$, $d = 0.31$, for both groups. No other significant effects were found for the P2.

All descriptive statistics of peak amplitudes and results of the separate Mixed Repeated Measure ANOVAs for N1 and P2 can be found in the **Supplementary Tables 1, 2**.

FIGURE 3 | Auditory event-related potentials [i.e., N1 (80–200 ms); P2 (200–300 ms)] for complex sounds in patients with schizophrenia (SQZ) and healthy controls (CTR): **(A)** for bird songs in the left hemisphere (i.e., cluster FC5) and **(B)** right hemisphere (i.e., cluster FC6); and **(C)** for human voices in the left hemisphere (i.e., cluster FC5) and **(D)** the right hemisphere (i.e., cluster FC6).

H2: Category—Complex Sounds (Voices vs. Bird Songs)

There were no main effects of stimulus category, nor category*group or category*group*hemisphere interactions on N1 amplitude (see also **Supplementary Table 3**).

A main effect of category was found on P2 amplitude, $F_{(1,51)} = 78.63$, $p < 0.001$, $\eta^2_p = 0.607$, indicating that voices elicited increased P2 amplitude (see **Figure 3**) than bird songs for both groups (all $p < 0.036$). No group or hemisphere*group interactions were found (see also **Supplementary Table 4**).

H3: Habituation Effects

For N1, we found a significant linear contrast effect of order for human voices, $F_{(1,51)} = 10.00$, $p = 0.003$, $\eta^2_p = 0.164$, but not for bird songs $F_{(1,51)} = 1.75$, $p = 0.192$, $\eta^2_p = 0.033$. No further effects were observed.

For P2, we found a significant linear contrast effect of order in bird songs $F_{(1,51)} = 21.71$, $p < 0.001$, $\eta^2_p = 0.299$, and voices, $F_{(1,51)} = 81.10$, $p < 0.001$, $\eta^2_p = 0.614$. Also, a significant interaction order*group was found in P2 for bird songs, $F_{(1,51)} = 4.21$, $p < 0.045$, $\eta^2_p = 0.760$ [vs. human voices, $F_{(1,51)} = 0.294$, $p = 0.590$, $\eta^2_p = 0.006$], independently of the hemisphere,
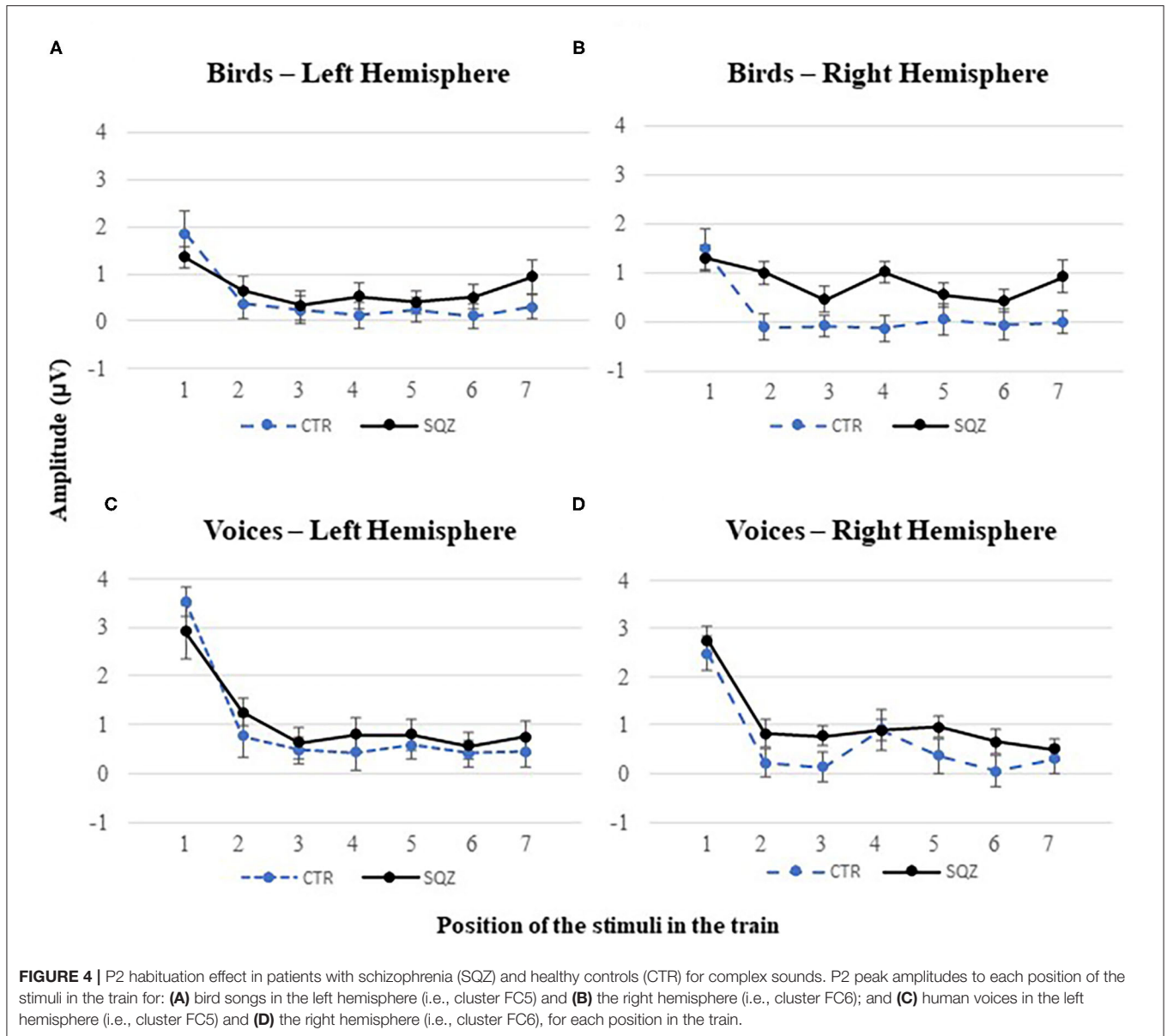
FIGURE 4 | P2 habituation effect in patients with schizophrenia (SQZ) and healthy controls (CTR) for complex sounds. P2 peak amplitudes to each position of the stimuli in the train for: **(A)** bird songs in the left hemisphere (i.e., cluster FC5) and **(B)** the right hemisphere (i.e., cluster FC6); and **(C)** human voices in the left hemisphere (i.e., cluster FC5) and **(D)** the right hemisphere (i.e., cluster FC6), for each position in the train.

$F_{(1,51)} = 1.02$, $p = 0.319$, $\eta^2_p = 0.020$. This finding (see **Figures 3, 4**) represents a gradual decrease in P2 amplitude for controls from Bird1 to Bird7 at both hemispheres (Bird1: $M_{FC5} = 1.85$, $SD = 2.44$, $M_{FC6} = 1.49$, $SD = 2.21$; Bird7: $M_{FC5} = 0.30$, $SD = 1.38$, $M_{FC6} = -0.01$, $SD = 1.31$, both $ps < 0.040$). This habituation effect on birds was not observed in patients (Bird1: $M_{FC5} = 1.35$, $SD = 1.19$, $M_{FC6} = 1.31$, $SD = 1.41$; Bird7: $M_{FC5} = 0.94$, $SD = 0.86$, $M_{FC6} = -0.92$, $SD = 1.75$, all $ps > 0.820$).

DISCUSSION

The habituation of the auditory N1 and P2 components of the ERPs remains largely unexplored in schizophrenia, especially when compared with the body of research directed to P50 and P300 components. The current work intended to expand the knowledge on these components. For this purpose, N1 and P2 amplitudes were measured in a roving paradigm. Our hypotheses

considered not only the habituation properties of N1 and P2, but also the different modulations induced by distinct stimulus categories. Supporting H1, we found a diminished N1 amplitude for pure tones in patients at the right frontotemporal region. In turn, H2 was not confirmed since human voices elicited higher P2 amplitudes than bird song at frontotemporal locations for both groups. Finally, H3 was partially verified, since habituation effects were moderated by group differences in P2, but not in N1 and only for bird songs. Our results showed that human voices (but not bird songs) specifically elicited a pattern of habituation in N1 in both groups. However, P2 amplitudes faded with train repetitions of bird songs only in the control group, while habituation effects to human voices were present in both groups.

The results concerning H1—Pure tones—resemble previous findings from oddball tasks, since patients with schizophrenia seem to have difficulties in allocating attention to target/relevant auditory stimuli, as reflected in the attenuation of the N1

component (9). However, for H2 we found no evidence that deficits in voice processing in schizophrenia would lead to a diminished P2. Notwithstanding, our results support previous literature that reported higher P2 amplitudes elicited by human voices compared to bird songs at Frontocentral locations (5, 7, 11), but these results did not interact with group. In other words, patients showed similar response amplitudes as controls in the initial discrimination of human voices, despite the literature describing abnormalities in P2 modulation for this group (24, 25). It may be the case that the P2 reflects an early preference for human voice processing that goes in line with the social nature of the human brain and the priority given to social cues within the species (6, 7). Although there have been descriptions of abnormalities regarding voice processing when negative emotional content is added (29), no differences have been found for non-semantic speech and neutral voice sounds (28, 30), suggesting that early stages of voice processing might be unaffected in early schizophrenia.

Taking together H1 and H2, N1 seems to mirror an orienting response toward target/simple stimuli that is impaired in schizophrenia, while P2 is likely to index more specific, elaborated categorization processes of complex sounds that are not necessarily impaired in schizophrenia. Given the differences in functional significance, these findings further unveil that the current experimental manipulation led to the dissociation of N1 and P2 amplitude, which is also reflected in habituation processes.

Furthermore, our results provide evidence that both N1 and P2 are habituating components (1, 13), notwithstanding that habituation effects are category dependent for N1. Interestingly, habituation effects were moderated by group differences in P2, but not in N1. This partially contradicts H3, as we expected a general dysfunction in habituation processes at both time-windows and stimuli in patients (25). Importantly, this effect was specific to birds and not voices, with patients showing the expected pattern of non-habituation for birds in P2. Patterns of abnormal habituation have been found for this component in schizophrenia using simple stimuli (25), and the same seems to be the case for bird songs [complex non-vocal stimuli. The specificity regarding auditory processing of the human voice may help to understand this differential effect. The early positivity that occurs around the P2 time-window has been described as sensitive to voices with increased amplitudes for these stimuli when compared with other sound categories (5). Our results corroborate these findings i.e., human voices elicited higher P2 amplitudes than bird songs for both groups]. In fact, the preference for sensory information of conspecifics is observable on the visual N170 and its increased amplitudes for faces when compared with objects (39, 40). Thus, it is possible that both auditory and visual preference for conspecific stimuli plays a relevant role in social information processing (41). From the available evidence on this matter, one study by Williams et al. (42) analyzed the habituation of brain activations in response to faces in schizophrenia. No habituation was found in several brain regions of patients (e.g., primary visual cortex and hippocampus), but when looking at the putative generator of the N170—the Fusiform Face Area (43)—patients and controls showed a similar pattern of habituation to faces. It remains plausible therefore

that conspecific signals (such as voice and faces) are processed by specialized brain modules that are somewhat independent of other stimulus categories, considering how crucial it is for communication and social interactions. Developmental studies highlight indeed that preference for voices emerges early, with habituation to human voices (vs. environmental stimuli) being observable in preschoolers (44). Additionally, brain regions responsible for voice processing are thought to mature during the 1st year of life (45), whereas P2 modulations by stimulus category and habituation are mature in 5-year-old toddlers (46, 47). So, it is possible that voice processing specializes much earlier than the onset of schizophrenia symptoms and possibly explaining how first-episode patients exhibit normal patterns of brain response to neutral non-speech human voices.

Limitations of this study comprise the absence of a formal assessment of mental and hearing abilities and the inclusion of groups on antipsychotic medication, making it impossible to dissociate the effects of the medication from the abnormalities found, as it has been previously suggested [e.g., (48)]. Also, only mid-latency ERP components were studied, not allowing us to investigate later stages of the auditory processing. Future research should consider additional stages of voice processing and investigate more nuanced associations with schizophrenia symptoms, such as hallucinations, as possible predictive factors of impairment [see e.g., (33)]. Nonetheless, regarding abnormal habituation for bird songs, it is important that future research uses other stimuli to understand if this pattern of habituation is generalized for non-voices or specific for bird songs. In addition, our results show that preference for human voices seems to be unaffected in early diagnosed schizophrenia patients. The investigation of abnormal patterns of brain activity in patients with schizophrenia, such as the ones found in the current work, allows researchers to better understand the impairments and symptoms of this pathology. More specifically, the ability to filter out irrelevant or repetitive information in early and later stages of information processing. This growing knowledge may be crucial for the future development of neurobiologically-informed early-diagnosis, assessment, and treatment for schizophrenia.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author, Prune Mazer, upon reasonable request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Local Ethics committee of the Faculty of Psychology and Educational Sciences of the University of Porto. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

FF-S, PA, TP, FB, and JM-T conceptualized the study: CC-R, TP, and CS collected the data: PM, IM, RP, FF-S, and TP analyzed

the data: PM, IM, RP, and TP drafted the manuscript. All authors revised the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyt.2021.630406/full#supplementary-material

REFERENCES

1. Özesmio Ç, Dolu N, Süer C, Gölgelio A, Aşçioglu M. Habituation of the auditory evoked potential in a short interstimulus interval paradigm. *Int J Neurosci.* (2000) 105:87–95. doi: 10.3109/00207450009003268

2. Pratt H. *Sensory ERP Components.* New York, NY: Oxford University Press (2011).

3. Luck SJ. *An Introduction to the Event-Related Potential Technique, 2nd edn.* Cambridge, MA; London, England: The MIT Press (2014).

4. Paiva TO, Almeida PR, Ferreira-Santos F, Vieira JB, Silveira C, Chaves PL, et al. Similar sound intensity dependence of the N1 and P2 components of the auditory ERP: averaged and single trial evidence. *Clin Neurophysiol.* (2016) 127:499–508. doi: 10.1016/j.clinph.2015.06.016

5. Charest I, Pernet CR, Rousselet GA, Quiñones I, Latinus M, Fillion-Bilodeau S, et al. Electrophysiological evidence for an early processing of human voices. *BMC Neurosci.* (2009) 10:127. doi: 10.1186/1471-2202-10-127

6. De Lucia M, Clarke S, Murray MM. A Temporal hierarchy for conspecific vocalization discrimination in humans. *J Neurosci.* (2010) 30:11210–21. doi: 10.1523/JNEUROSCI.2239-10.2010

7. Capilla A, Belin P, Gross J. The early spatio-temporal correlates and task independence of cerebral voice processing studied with MEG. *Cereb Cortex.* (2013) 23:1388–95. doi: 10.1093/cercor/bhs119

8. Schirmer A, Gunter TC. Temporal signatures of processing voiceness and emotion in sound. *Soc Cogn Affect Neurosci.* (2017) 12:902–9. doi: 10.1093/scan/nsx020

9. Rosburg T, Boutros NN, Ford JM. Reduced auditory evoked potential component N100 in schizophrenia - a critical review. *Psychiatry Res.* (2008) 161:259–74. doi: 10.1016/j.psychres.2008.03.017

10. Tomé D, Sampaio M, Mendes-Ribeiro J, Barbosa F, Marques-Teixeira J. Auditory event-related potentials in children with benign epilepsy with centro-temporal spikes. *Epilepsy Res.* (2014) 108:1945–9. doi: 10.1016/j.eplepsyres.2014.09.021

11. Bruneau N, Roux S, Cléry H, Rogier O, Bidet-Caulet A, Barthélémy C. Early neurophysiological correlates of vocal versus non-vocal sound processing in adults. *Brain Res.* (2013) 1528:20–7. doi: 10.1016/j.brainres.2013.06.008

12. Rentzsch J, Jockers-Scherübl M, Boutros N, Gallinat J. Test-retest reliability of P50, N100 and P200 auditory sensory gating in healthy subjects. *Int J Psychophysiol.* (2008) 67:81–90. doi: 10.1016/j.ijpsycho.2007.10.006

13. Pereira DR, Cardoso S, Ferreira-Santos F, Fernandes C, Cunha-Reis C, Paiva TO, et al. Effects of inter-stimulus interval (ISI) duration on the N1 and P2 components of the auditory event-related potential. *Int J Psychophysiol.* (2014) 94:311–8. doi: 10.1016/j.ijpsycho.2014.09.012

14. Bramon E. Meta-analysis of the P300 and P50 waveforms in schizophrenia. *Schizophr Res.* (2004) 70:315–29. doi: 10.1016/j.schres.2004.01.004

15. Atagun MI, Drukker M, Hall MH, Altun IK, Tatli SZ, Guloksuz S, et al. Meta-analysis of auditory P50 sensory gating in schizophrenia and bipolar disorder. *Psychiatry Res Neuroimaging.* (2020) 300:111078. doi: 10.1016/j.pscychresns.2020.111078

16. McCarley RW, Faux SF, Shenton ME, Nestor PG, Adams J. Event-related potentials in schizophrenia: their biological and clinical correlates and new model of schizophrenic pathophysiology. *Schizophr Res.* (1991) 4:209–31. doi: 10.1016/0920-9964(91)90034-O

17. Chang W-P, Arfken CL, Sangal MP, Boutros NN. Probing the relative contribution of the first and second responses to sensory gating indices: a meta-analysis: S1, S2, and the gating ratio. *Psychophysiology.* (2011) 48:980–92. doi: 10.1111/j.1469-8986.2010.01168.x

18. Force RB, Venables NC, Sponheim SR. An auditory processing abnormality specific to liability for schizophrenia. *Schizophr Res.* (2008) 103:298–310. doi: 10.1016/j.schres.2008.04.038

19. Salisbury DF, Collins KC, McCarley RW. Reductions in the N1 and P2 Auditory Event-Related Potentials in First-Hospitalized and Chronic Schizophrenia. *Schizophr Bull.* (2010) 36:991–1000. doi: 10.1093/schbul/sbp003

20. Foxe JJ, Yeap S, Snyder AC, Kelly SP, Thakore JH, Molholm S. The N1 auditory evoked potential component as an endophenotype for schizophrenia: high-density electrical mapping in clinically unaffected first-degree relatives, first-episode, and chronic schizophrenia patients. *Eur Arch Psychiatry Clin Neurosci.* (2011) 261:331–9. doi: 10.1007/s00406-010-0176-0

21. del Re EC, Spencer KM, Oribe N, Mesholam-Gately RI, Goldstein J, Shenton ME, et al. Clinical high risk and first episode schizophrenia: auditory event-related potentials. *Psychiatry Res Neuroimaging.* (2015) 231:126–33. doi: 10.1016/j.pscychresns.2014.11.012

22. Wu K-Y, Chao C-W, Hung C-I, Chen W-H, Chen Y-T, Liang S-F. Functional abnormalities in the cortical processing of sound complexity and musical consonance in schizophrenia: evidence from an evoked potential study. *BMC Psychiatry.* (2013) 13:158. doi: 10.1186/1471-244X-13-158

23. Mathalon DH, Roach BJ, Ferri JM, Loewy RL, Stuart BK, Perez VB, et al. Deficient auditory predictive coding during vocalization in the psychosis risk syndrome and in early illness schizophrenia: the final expanded sample. *Psychol Med.* (2019) 49:1897–904. doi: 10.1017/S0033291718002659

24. Ferreira-Santos F, Silveira C, Almeida PR, Palha A, Barbosa F, Marques-Teixeira J. The auditory P200 is both increased and reduced in schizophrenia? A meta-analytic dissociation of the effect for standard and target stimuli in the oddball task. *Clin Neurophysiol.* (2012) 123:1300–8. doi: 10.1016/j.clinph.2011.11.036

25. Boutros NN, Korzyukov O, Jansen B, Feingold A, Bell M. Sensory gating deficits during the mid-latency phase of information processing in medicated schizophrenia patients. *Psychiatry Res.* (2004) 126:203–15. doi: 10.1016/j.psychres.2004.01.007

26. Morales-Muñoz I, Jurado-Barba R, Fernández-Guinea S, Rodríguez-Jiménez R, Jiménez-Arriero MÁ, Criado JR, et al. Sensory gating deficits in first-episode psychosis: evidence from neurophysiology, psychophysiology, and neuropsychology. *J Nerv Ment Dis.* (2016) 204:877–84. doi: 10.1097/NMD.0000000000000572

27. Rosburg T. Auditory N100 gating in patients with schizophrenia: a systematic meta-analysis. *Clin Neurophysiol.* (2018) 129:2099–111. doi: 10.1016/j.clinph.2018.07.012

28. Pinheiro AP, del Re E, Mezin J, Nestor PG, Rauber A, McCarley RW, et al. Sensory-based and higher-order operations contribute to abnormal emotional prosody processing in schizophrenia: an electrophysiological investigation. *Psychol. Med.* (2013) 43:603–18. doi: 10.1017/S003329171200133X

29. Pinheiro AP, Rezaii N, Rauber A, Liu T, Nestor PG, McCarley RW, et al. Abnormalities in the processing of emotional prosody from single words in schizophrenia. *Schizophr Res.* (2014) 152:235–41. doi: 10.1016/j.schres.2013.10.042

30. Liu T, Pinheiro AP, Zhao Z, Nestor PG, McCarley RW, Niznikiewicz M. Simultaneous face and voice processing in schizophrenia. *Behav Brain Res.* (2016) 305:76–86. doi: 10.1016/j.bbr.2016.01.039

31. Hsieh MH, Lin Y-T, Chien Y-L, Hwang T-J, Hwu H-G, Liu C-M, et al. Auditory event-related potentials in antipsychotic-free subjects with ultra-high-risk state and first-episode psychosis. *Front Psychiatry.* (2019) 10:223. doi: 10.3389/fpsyt.2019.00223

32. Fuerst DR, Gallinat J, Boutros NN. Range of sensory gating values and test?retest reliability in normal subjects. *Psychophysiology.* (2007) 44:620–6. doi: 10.1111/j.1469-8986.2007.00524.x

33. Conde T, Gonçalves OF, Pinheiro AP. A cognitive neuroscience view of voice-processing abnormalities in schizophrenia: a window into auditory verbal hallucinations? *Harv Rev Psychiatry.* (2016) 24:148–63. doi: 10.1097/HRP.0000000000000082

34. de Wilde OM, Bour LJ, Dingemans PM, Koelman JHTM, Linszen DH. A meta-analysis of P50 studies in patients with schizophrenia and relatives: differences in methodology between research groups. *Schizophr Res.* (2007) 97:137–51. doi: 10.1016/j.schres.2007.04.028

35. Baldeweg T, Klugman A, Gruzelier J, Hirsch SR. Mismatch negativity potentials and cognitive impairment in schizophrenia. *Schizophr Res.* (2004) 69:204–17. doi: 10.1016/j.schres.2003.09.009

36. Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *J Neurosci Methods.* (2004) 134:9–21. doi: 10.1016/j.jneumeth.2003.10.009

37. Delorme A, Mullen T, Kothe C, Akalin Acar Z, Bigdely-Shamlo N, Vankov A, et al. EEGLAB, SIFT, NFT, BCILAB, and ERICA: new tools for advanced EEG processing. *Comput Intell Neurosci.* (2011) 2011:130714. doi: 10.1155/2011/130714

38. Viola FC, Thorne J, Edmonds B, Schneider T, Eichele T, Debener S. Semi-automatic identification of independent components representing EEG artifacts. *Clin Neurophysiol.* (2009) 120:868–77. doi: 10.1016/j.clinph.2009.01.015

39. Mercure E, Kadosh KC, Johnson MH. The N170 shows differential repetition effects for faces, objects, and orthographic stimuli. *Front Hum Neurosci.* (2011) 5:6. doi: 10.3389/fnhum.2011.00006

40. Hinojosa JA, Mercado F, Carretié L. N170 sensitivity to facial expression: a meta-analysis. *Neurosci Biobehav Rev.* (2015) 55:498–509. doi: 10.1016/j.neubiorev.2015.06.002

41. Liu T, Pinheiro A, Zhao Z, Nestor PG, McCarley RW, Niznikiewicz MA. Emotional cues during simultaneous face and voice processing: electrophysiological insights. *PLoS One.* (2012) 7:e31001. doi: 10.1371/journal.pone.0031001

42. Williams LE, Blackford JU, Luksik A, Gauthier I, Heckers S. Reduced habituation in patients with schizophrenia. *Schizophr Res.* (2013) 151:124–32. doi: 10.1016/j.schres.2013.10.017

43. Kanwisher N, Yovel G. The fusiform face area: a cortical region specialized for the perception of faces. *Philos Trans R Soc B Biol Sci.* (2006) 361:2109–28. doi: 10.1098/rstb.2006.1934

44. Yamasaki T, Ogata K, Maekawa T, Ijichi I, Katagiri M, Mitsudo T, et al. Rapid maturation of voice and linguistic processing systems in preschool children: a near-infrared spectroscopic study. *Exp Neurol.* (2013) 250:313–20. doi: 10.1016/j.expneurol.2013.10.005

45. Dubois J, Poupon C, Thirion B, Simonnet H, Kulikova S, Leroy F, et al. Exploring the early organization and maturation of linguistic pathways in the human infant brain. *Cereb Cortex.* (2016) 26:2283–98. doi: 10.1093/cercor/bhv082

46. Ponton CW, Eggermont JJ, Kwong B, Don M. Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. *Clin Neurophysiol.* (2000) 111:220–36. doi: 10.1016/S1388-2457(99)00236-9

47. Mahajan Y, McArthur G. Maturation of auditory event-related potentials across adolescence. *Hear Res.* (2012) 294:82–94. doi: 10.1016/j.heares.2012.10.005

48. Valkonen-Korhonen M, Purhonen M, Tarkka IM, Sipilä P, Partanen J, Karhu J, et al. Altered auditory processing in acutely psychotic never-medicated first-episode patients. *Cogn Brain Res.* (2003) 17:747–58. doi: 10.1016/S0926-6410(03)00199-X