Check for updates

Hypersensitivity to passive voice hearing in hallucination proneness

Joseph F. Johnson[1], Michel Belyk[2], Michael Schwartze[1],
Ana P. Pinheiro[3] and Sonja A. Kotz[1,4]*

[1]Department of Neuropsychology and Psychopharmacology, University of Maastricht, Maastricht,
Netherlands, [2]Department of Psychology, Edge Hill University, Ormskirk, United Kingdom,
[3]Faculdade de Psicologia, Universidade de Lisboa, Lisbon, Portugal, [4]Department
of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences,
Leipzig, Germany

Voices are a complex and rich acoustic signal processed in an extensive cortical brain network. Specialized regions within this network support voice perception and production and may be differentially affected in pathological voice processing. For example, the experience of hallucinating voices has been linked to hyperactivity in temporal and extra-temporal voice areas, possibly extending into regions associated with vocalization. Predominant self-monitoring hypotheses ascribe a primary role of voice production regions to auditory verbal hallucinations (AVH). Alternative postulations view a generalized perceptual salience bias as causal to AVH. These theories are not mutually exclusive as both ascribe the emergence and phenomenology of AVH to unbalanced top-down and bottom-up signal processing. The focus of the current study was to investigate the neurocognitive mechanisms underlying predisposition brain states for emergent hallucinations, detached from the effects of inner speech. Using the temporal voice area (TVA) localizer task, we explored putative hypersalient responses to passively presented sounds in relation to hallucination proneness (HP). Furthermore, to avoid confounds commonly found in in clinical samples, we employed the Launay-Slade Hallucination Scale (LSHS) for the quantification of HP levels in healthy people across an experiential continuum spanning the general population. We report increased activation in the right posterior superior temporal gyrus (pSTG) during the perception of voice features that positively correlates with increased HP scores. In line with prior results, we propose that this right-lateralized pSTG activation might indicate early hypersensitivity to acoustic features coding speaker identity that extends beyond own voice production to perception in healthy participants prone to experience AVH.

KEYWORDS

temporal voice area (TVA), voice perception, hallucination proneness, functional magnetic brain imaging (fMRI), neuroimaging, salience account

Introduction

The human voice is a complex signal that carries rich information. This allows the listener not only to identify linguistic messages but also who speaks and how something is said (Belin et al., 2004; Lavan et al., 2019). Some individuals experience auditory verbal hallucinations (AVH), in which they perceive voices in the absence of a corresponding incoming voice signal (Bentall, 1990; Anthony, 2004; Brookwell et al., 2013). Experience of AVH is a key symptom of schizophrenia (Bauer et al., 2011; Larøi et al., 2012; Hugdahl and Sommer, 2018). Yet, it is also reported in multiple other psychiatric, developmental, and neurological disorders (Van Os et al., 2000; Reininghaus et al., 2016; Waters and Fernyhough, 2017; Rollins et al., 2019; Zhuo et al., 2019) and in a minority of otherwise healthy people (Beavan et al., 2011; Linscott and Van Os, 2013; McGrath et al., 2015). Variability in AVH phenomenology exists within and across brain disorders (Stephane et al., 2003; Jones, 2010) and between clinical and non-clinical voice hearers (Daalman et al., 2011; Larøi et al., 2012; Johns et al., 2014; Baumeister et al., 2017). However, hallucinated voices commonly carry information regarding the identity or emotion of a perceived speaker (Stephane et al., 2003; Larøi and Woodward, 2007; Badcock and Chhabra, 2013; McCarthy-Jones et al., 2014), therefore involving a wide range of cortical areas in a voice perception network. Multiple cognitive theories have been proposed delineating the emergence and phenomenology of AVH (Jones, 2010; Ćurčić-Blake et al., 2017; Rollins et al., 2019). One long standing model considers hallucinations as the misattribution of self-generated input to an outside source (Feinberg, 1978). In terms of AVH, signals from voice production cortical regions during inner speech are misperceived as hearing someone else speak (Allen et al., 2007a; Jones and Fernyhough, 2007a,b; Swiney and Sousa, 2014; Gregory, 2016). Recently, competing theories have gained traction, claiming that the initiation of hallucinations does not require motor activity while they are, at their core, misperceived sensations from the environment (e.g., Ford and Mathalon, 2019; Thakkar et al., 2021).

The selection and processing of sensory inputs from the environment relevant to learning, adaptation, or behavioral responses involves multiple regions and distributed networks across the brain. The role of salience attribution within this integrated system provides the necessary trigger to shift processing from a state of rest to active sensation and perception (Menon and Uddin, 2010; Menon, 2011; Palaniyappan and Liddle, 2012; Uddin, 2015). According to this framework, increased auditory cortex activation associated with AVH can be ascribed to a bottom-up hypersensitivity, or salience bias, toward irrelevant sounds. The modulation and over-weighting of top-down predictions may influence this salience bias as well as guide the system to perceive what it expects in meaningless unimodal and multimodal stimuli (Friston,

2005, 2012; Fletcher and Frith, 2009; Deneve and Jardri, 2016; Jardri et al., 2016; Leptourgos et al., 2017). Since voice signals in humans are inherently salient to human listeners, they may be particularly implicated in hypersensitive responses leading to false perceptions. Furthermore, for those who experience AVH, the engagement of brain regions controlling inner speech signals, memory retrieval, and emotion may then guide the phenomenology of the perceived speech in terms of content and speaker-related features (Waters et al., 2012). Abnormal salience processing has been strongly linked to positive symptoms of schizophrenia (Miyata, 2019).

Researching the contribution of these mechanisms to AVH in non-clinical samples may be particularly useful as it avoids potential confounds seen in clinical populations such as medication, age of onset, and duration of symptoms that may affect brain structure and function (Verdoux and van Os, 2002; Kelleher et al., 2010; Kelleher and Cannon, 2011). This perspective is in line with the experiential continuum of psychosis (Johns, 2005; Beavan et al., 2011; Larøi et al., 2012; de Leede-Smith and Barkus, 2013; Johns et al., 2014; Zhuo et al., 2019), whereby functional variability in the mechanisms serving perception across the population account for the spectrum of normal experience, vivid perceptions and imagery, sub-clinical forms of hallucinations, and those seen in full-blown psychosis. The revised Launay-Slade Hallucination Scale (LSHS) is as a measure of perceptual experience and beliefs associated with vivid daydreams, thoughts, imagery, and those related to false perceptions such as visual and auditory hallucinations (Larøi and Van Der Linden, 2005). The LSHS provides a measure of hallucination proneness (HP), where higher scores signify increasing abnormality in perceptual experience and beliefs, including true hallucinations. Although individual items from the LSHS can be used to identify the prevalence of AVH (e.g., Kompus et al., 2015), HP itself is not a measure of risk for psychosis.

Two critical factors have been incorporated into the formulation of our hypotheses. First, differential brain activity may indicate abnormal voice processing as a predisposition for false perceptions, i.e., activation patterns similar to those during hallucinations. Second, the localization of reported changes in brain responses may indicate a specific stage within hierarchical voice processing at which this predisposition manifests. To date, no consensus has been empirically established regarding a trait-based association between hallucinations and brain responses to the voice. For example, when presented with voices, patients who commonly experience hallucinations display decreases (Copolov et al., 2003), increases (Martí-Bonmatí et al., 2007; Parellada et al., 2008; Escartí et al., 2010), or no activation differences in voice selective temporal regions (Woodruff et al., 1997; Simons et al., 2010). Such inconsistency is likely due to methodological

heterogeneity (Bohlken et al., 2017). For example, these studies differed in terms of stimulus type, stimulus content, and the inclusion of a non-hallucinating patient control group. Moreover, patients with chronic hallucinations can experience spontaneous AVH during scanning (Jardri et al., 2011; Kühn and Gallinat, 2012; van Lutterveld et al., 2013; Zmigrod et al., 2016), which may even be unintentionally elicited by tasks (e.g., Copolov et al., 2003; Parellada et al., 2008). Although this hallucinatory state elicits brain activity in voice perception regions, simultaneous external voice input during AVH results in a paradoxical net activity decrease (Kompus et al., 2011; Hugdahl and Sommer, 2018).

The localization of changes in functional brain activity within the voice processing network can be particularly informative in determining how HP may arise. Within the upper bank and lateral regions of the temporal lobe, voice signals are processed hierarchically along a pathway composed of multiple functional subsystems or components (Belin et al., 2004; Pernet et al., 2015; Zhang et al., 2021). The engagement of these temporal voice areas (TVA) starts with the evaluation of low-level acoustic features in the posterior superior temporal gyrus (STG), an area specialized in processing spectro-temporal properties of complex sounds (Griffiths and Warren, 2004; Warren J. D. et al., 2005; Warren J. E. et al., 2005). Further processing occurs along hemispherically specialized pathways, with linguistic features predominantly in the left and paralinguistic (i.e., speaker-related information) in the right side of the brain (Belin et al., 2000; Formisano et al., 2008). However, some stimuli such as emotional vocalizations contain both speaker-and speech-relevant information and involve bilateral processing of separate features in the signal (Schirmer and Kotz, 2006). Importantly, AVH often contain marked paralinguistic information about speaker identity or emotion (Larøi and Woodward, 2007; Larøi et al., 2012; McCarthy-Jones et al., 2014). In non-clinical voice hearers, however, the degree of perceived emotional valence is less prominent (Daalman et al., 2012; de Boer et al., 2016). Speaker-related feature processing operates along a multi-stage hierarchy in the right temporal cortex along a posterior to anterior gradient (Nakamura et al., 2001; Belin and Zatorre, 2003; von Kriegstein et al., 2003; von Kriegstein and Giraud, 2004). The TVA localizer is a widely used fMRI task which reliably identifies activation peaks localized in the bilateral anterior, middle, and posterior superior temporal cortex (Pernet et al., 2015). By comparing voice to non-voice activation in response to passively heard sounds, regions of interest (ROI) can be defined for further investigation. Using ROIs produced by this task, we predicted HP-related early sensitivity to low-level voice features to be isolated to the posterior STG ROI. Alternatively, changes to voice processing in the anterior direction of the right STG might indicate an abnormal salience bias for identity or emotion associated with an increasing propensity to hallucinate.

Methods

Participants

Twenty-six participants took part in this study, recruited through the SONA system and social media channels at Maastricht University, Netherlands. Participants were provided with informed consent and offered university study credit for compensation. Exclusion criteria included any history of psychotic disorder, neurological impairment, history of drug dependence or abuse, and traumatic brain injury. Participants were screened for MRI safety and reported no metal implants, claustrophobia, or pregnancy. Furthermore, all participants reported no known hearing deficits. Robust statistics using the interquartile range rule for participant age revealed one outlier (Rousseeuw and Hubert, 2011), leading to the exclusion of the dataset from further analysis. Of the resulting 25 individuals (17 female), the average age was 20.92 years (SD 3.95; range 18–32). The Ethical Review Committee of the Faculty of Psychology and Neuroscience at Maastricht University (ERCPN-176_08_02_2017) approved this study.

Hallucination proneness

The revised LSHS was employed as a self-report measure of HP (Larøi and Van Der Linden, 2005). The questionnaire consists of 16 items targeting tactile, sleep-related, visual, and auditory modalities of psychosis-like experience as well as vivid thoughts and daydreaming. Responses were given using a five-point Likert scale, measuring the extent to which each statement applied to them. The sum of all responses equated to an overall HP measure. Furthermore, to investigate the exclusivity of auditory-only items, subscores of three items were summed to produce a composite score (Larøi et al., 2004; Larøi and Van Der Linden, 2005).

Voice area fMRI-localizer task

Voice selective cortical brain regions were identified using a standard fMRI-localizer task (Belin et al., 2000). This widely used tool reliably probes activity across three bilateral peaks in the superior temporal gyrus (e.g., Pernet et al., 2015), often designated as anterior, middle, and posterior temporal voice areas (TVA). Furthermore, many studies applying this task have reported extra-temporal voice regions, such as the inferior frontal cortex (IFC). The voice area localizer consists of 20 vocal (V) and 20 non-vocal (NV) trials. Additionally, 20 silence (S) trials are included allowing relaxation of the hemodynamic response to auditory stimuli. The voice condition is composed of human speech (words, syllables, or sentence excerpts) and non-speech voices produced by male and female speakers of different

ages (7 babies, 12 adults, 23 children, and 5 elderly). This broad selection of voice stimuli allows for the probing and inclusion of functionally diverse regions of TVA. Conversely, the non-voice condition includes environmental (natural and animal) and man-made (e.g., cars, alarm clocks, instrumental music) sounds. Sound clips are presented at a standard 70 db volume (for a detailed report of the included sounds and recording duration, amplitude, and frequency see Pernet et al., 2015). Trials were presented in a pseudorandom order, each with a duration of eight seconds. With a two second inter-trial interval, the total run time of the task was 10 min.

FMRI data acquisition

Scanning was conducted using a Siemens 3T Magnetom Prisma Fit equipped with a 32-channel head coil (Siemens Healthcare, Erlangen, Germany), at the Scannexus facilities (Maastricht, Netherlands). Structural whole-brain T1-weighted images were acquired with a single-shot echoplanar imaging (EPI) sequence [field of view (FOV) 256 mm; 192 axial slices; 1 mm slice thickness; 1 mm × 1 mm × 1 mm voxel size; repetition time (TR) of 2250 ms; echo-time (TE) 2.21 ms]. For the functional localizer task, T2-weighted EPI scans were collected (FOV 208 mm; 60 axial slices; 2 mm slice thickness; 2 mm × 2 mm × 2 mm voxel size; TE 30 ms; flip angle = 77°). To reduce scanner noise interference, auditory stimuli were presented *via* S14 MR-compatible earphones, fitted with foam earplugs (Sensimetrics Corporation). Furthermore, to provide relative silence during playback of auditory stimuli, a long inter-acquisition-interval was adopted where time between consecutive acquisition was delayed, resulting in a TR of 10 s. The delayed TR was timed to allow a 2,000 ms acquisition period during peak activation in the auditory cortex (Belin et al., 1999; Hall et al., 1999).

Data pre-processing and analysis

Pre-processing of the TVA localizer blood-oxygen-level-dependent (BOLD) signal was conducted in SPM12 (Wellcome Department of Cognitive Neurology, London, United Kingdom). A standard pipeline was applied using slice timing correction, realignment and unwarping, segmentation, normalization to standard (MNI) space (Fonov et al., 2009), and 8 mm isotropic Gaussian kernel full width at half maximum (FWHM) smoothing. Analysis followed a two-level procedure in which contrast estimates were first determined as fixed effects at the level of individual participants then modeled as random effects at the level of the sample. Contrast estimates were computed on BOLD data to assess voice sensitivity (V > NV) and sensitivity to environmental sounds (NV > S) for each participant. A first-level fixed-effects GLM analysis for the

conjunction analysis [(V > NV) ∩ (V > S)] was computed to localize the temporal voice areas. A second-level random-effects analysis tested for group-level significance and determined the ROIs for parameter extraction. Contrast estimates of V > S and NV > S were then used to contrast voice with non-voice activity, corrected for baseline, in the subsequent hypothesis-driven ROI analysis to investigate the correlation of voice-preferential TVA activity compared to HP. Contrast estimates were extracted from a 5 mm radius of the center coordinates from each region of peak activity produced in the TVA-localizer using the SPM MARSbar toolbox (Brett et al., 2002). Pearson's correlation analysis using bootstrapping (5000 samples) and bias-corrected confidence intervals was then employed to test for significant relationships between the sensitivity of the voice ROIs and HP measures.

Results

Hallucination proneness

For the HP composite score (possible maximum score of 80), the mean self-reported rating was 25.20 (SD 10.47; range 0–42). The HP auditory subscale mean score (possible maximum score 15) was 3.92 (SD 2.74; range 0–11). To test for normality of the distribution of demographics and HP across the sample, Shapiro–Wilk tests were conducted. Both total LSHS (0.948, df = 25, p = 0.229) and auditory subscale (0.928, df = 25, p = 0.078) were not different from normal. A moderately strong correlation was also found between LSHS auditory subscale and non-auditory item totals (r = 0.457, df = 25, p = 0.019).

Voice area localizer

The fMRI localizer task produced 5 clusters covering bilateral lateral temporal cortices, bilateral inferior frontal gyri, and the right precentral gyrus (preCG) (**Table 1** and **Figure 1**). Within each bilateral temporal cortex "voice patch," peak activity localizations were distinguished in three distinct regions: posterior (pSTG), middle (mSTG), and anterior STG (aSTG). These regions correspond to the expected divisions of the TVA localizer (Pernet et al., 2015).

FMRI correlation

Correlational tests were performed between contrast estimates representing voice preference [(V > S) > (NV > S)] observed in each TVA-ROI with both the composite HP score and the auditory subscore of the LSHS. All thresholds for significance were Bonferroni-adjusted for multiple comparisons using (p < 0.025). Only the right pSTG reached statistical
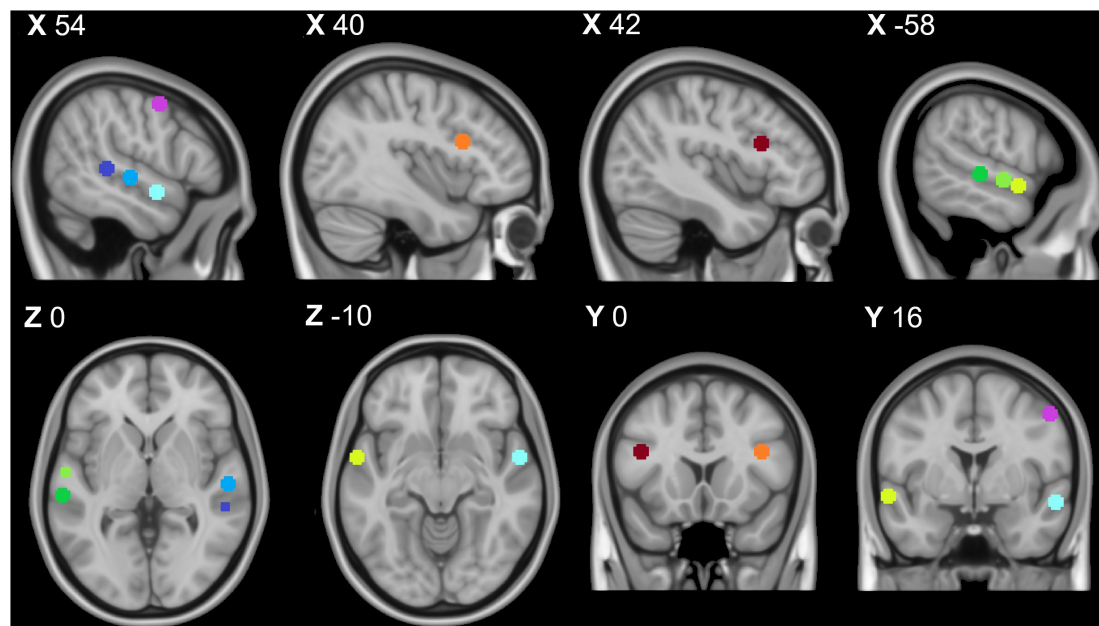
FIGURE 1

Temporal voice area fMRI localizer task results: Purple = right premotor cortex, dark blue = right posterior temporal gyrus, middle blue = right middle temporal gyrus, light blue = right anterior temporal gyrus, orange = right inferior frontal cortex, dark green = left posterior superior temporal gyrus, middle green = left middle superior temporal gyrus, light green = left anterior superior temporal gyrus, and red = left inferior temporal cortex. All coordinates listed in MNI space (x,y,z). This image was created using the FSL toolbox fsleyes (McCarthy, 2022).

TABLE 1 Results from temporal voice area fMRI localizer task.

Cluster #	Hem.	Label	BA	x	y	z	Cluster-Level p-FDR	Peak-Level p-FDR	Size (voxels)
1	L	mSTG	22	−58	−10	−4	1.6782E-17	1.4637E-09	4145
		pSTG	22	−60	−26	0		1.4637E-09	
		aSTG	22	−58	0	−8		1.3575E-08	
2	R	mSTG	22	56	−18	−2	2.0689E-17	1.4637E-09	4010
		aSTG	22	56	0	−12		1.6043E-08	
		pSTG	22	54	−34	4		1.6043E-08	
3	R	pMC	6	52	2	48	0.0049	4.1457E-05	285
4	L	IFC	44	−42	16	22	0.0383	0.0018	142
5	R	IFC	44	40	16	22	0.0227	0.0302	180

Hem, hemisphere; (a/m/p) STG, (anterior/middle/posterior) superior temporal gyrus; pMC, premotor cortex; IFC, inferior frontal cortex; BA, Brodmann's Area; p-FDR, false discovery rate corrected p-value (threshold = 0.05). All coordinates listed in MNI space (x, y, z).

significance ($r = 0.470$, df = 25, $p = 0.020$) (**Table 2** and **Figure 2**). *Post hoc* correlation analyses were run to assess the relative contributions of both voice (V > S) and non-voice (NV > S) contrasts to correlational analyses (see detailed results in **Supplementary Material**). We conducted these analyses in order to rule out a general hypersensitivity of temporal cortex activity non-specific to the conditions of interest probed by the conjunction analysis. No significant correlations with HP were found in any ROI for voice (V > S), however, a significant negative correlation was reported in the right IFC for non-voice (V > S) sensitive activity ($r = -0.614$, df = 25, $p = 0.001$).

Discussion

The current study investigated whether a measure of abnormal perceptual experience (HP) in a non-clinical sample is associated with variability in the functional brain responses of the temporal cortex regions serving detecting and processing of voice signals. Considering the well-established roles of specific voice sensitive regions of the cerebral cortex, we aimed to determine if this putative relationship would be limited to specific subprocesses in hierarchical voice perception. As hypothesized, activity for voice versus non-voice processing correlated positively with HP only in the pSTG, a region

TABLE 2 Voice preference response [(Voice > Silence) > (Non-voice > Silence)] correlation with hallucination proneness results.

Hem.	Label	ROI			LSHS		LSHS-Auditory	
		μ	SD	CI (95%)	r	p	r	p
L	aSTG	1.189	0.479	0.203–0.434	0.120	0.576	0.178	0.406
	mSTG	1.505	0.586	0.997–1.380	−0.237	0.267	−0.024	0.915
	pSTG	1.511	0.560	1.271–1.740	−0.058	0.791	0.055	0.797
R	aSTG	1.019	0.452	0.838–1.200	0.266	0.208	0.165	0.440
	mSTG	1.295	0.515	1.089–1.501	−0.177	0.408	−0.033	0.882
	pSTG	1.213	0.406	1.051–1.375	0.470	*0.020	0.276	0.192
R	pMC	0.625	0.447	0.446–0.804	0.087	0.685	−0.103	0.635
L	IFC	0.319	0.288	0.204–0.434	−0.048	0.827	−0.025	0.911
R	IFC	0.293	0.323	0.164–0.422	0.231	0.277	0.134	0.534

ROI, region of interest; (a/m/p) STG, (anterior/middle/posterior) superior temporal gyrus; pMC, premotor cortex; IFC, inferior frontal cortex; μ, mean activation from contrast; SD, standard deviation; LSHS, Launay-Slade Hallucination Proneness scale; LSHS-Auditory, subset of 3 auditory items, r = correlation coefficient, Bonferroni-corrected significance level (*$p < 0.025$).

FIGURE 2
Hallucination proneness fMRI correlation analysis results: Right posterior superior temporal gyrus (BA 22; MNI 54, −34, 4), Voice preference = contrast estimate [(Voice > Silence) > (Non-voice > Silence)], LSHS = Launay Slade Hallucination Proneness scale. Correlation coefficient $r = 0.470$, $df = 25$, $p = 0.020$.
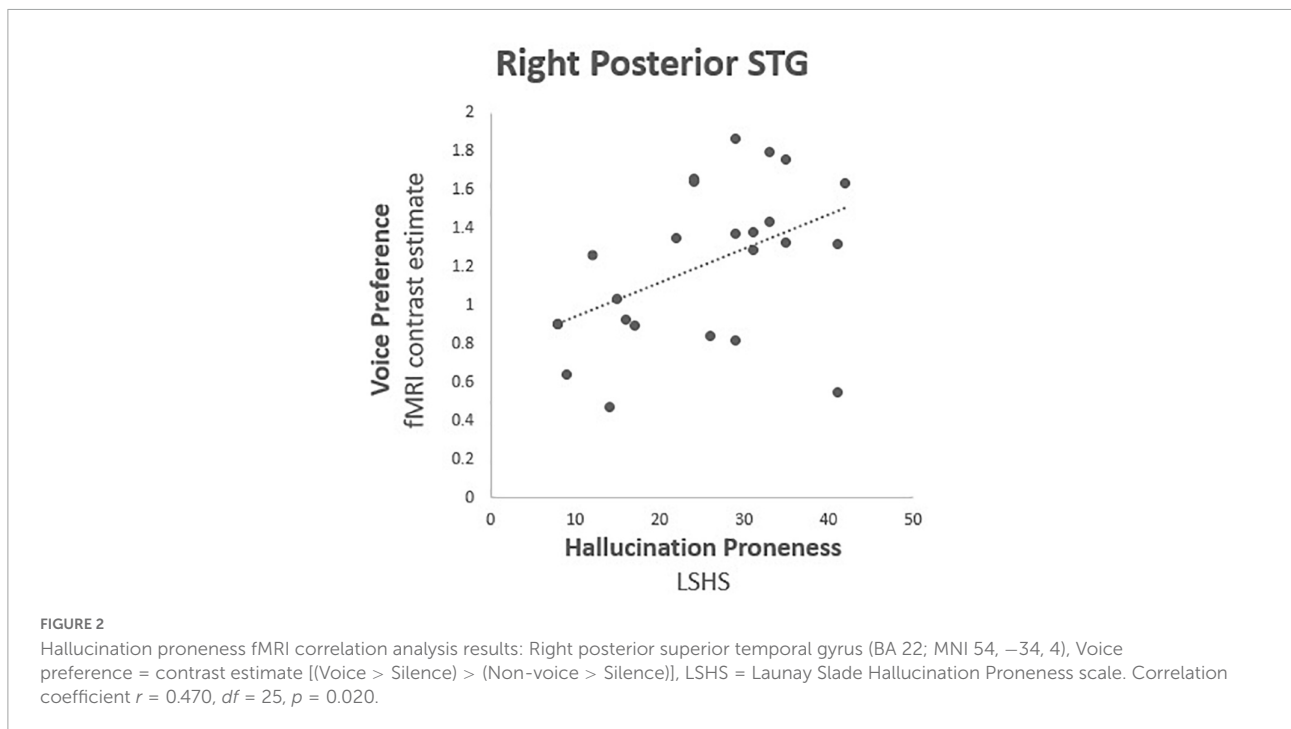
associated with the early processing of low-level acoustic features in complex auditory signals (i.e., Griffiths and Warren, 2004; Warren J. D. et al., 2005; Warren J. E. et al., 2005). Furthermore, this finding was restricted to the right hemisphere and therefore is likely linked to the processing of paralinguistic voice information (Belin et al., 2000; Formisano et al., 2008). Additionally, *post hoc* analysis revealed a negative correlation with HP in the right IFC for non-voice versus silence. Together, these findings may confirm that as the propensity to hallucinate increases, right posterior temporal lobe voice hypersensitivity increases and is accompanied by a decreased prefrontal response to non-vocal environmental sounds.

Hallucination proneness and hypersensitivity

Multiple neurocognitive mechanisms underlying hallucinations have been proposed. Most commonly, these theories have focused on describing the emergence and phenomenology of pathological voice hearing in patients with psychotic disorders such as schizophrenia (Allen et al., 2008; Hugdahl, 2015; Ćurčić-Blake et al., 2017). The most influential models describe atypical increases in brain activity in cortical voice regions. The current investigation was approached from the perspective of perceptual salience models claiming a central role of hypersensitivity to irrelevant sensory stimuli in auditory

regions (Menon and Uddin, 2010; Menon, 2011; Palaniyappan and Liddle, 2012; Uddin, 2015). Conversely, prominent self-monitoring models of hallucinatory experience describe increased activity as the result of insufficient suppression of sensory cortices during inner speech (Frith and Done, 1988; Weiss and Heckers, 1999; Tracy and Shergill, 2006; Allen et al., 2007a, 2008; Jones and Fernyhough, 2007b). According to this theory, the activation of speech production regions is required for the emergence of AVH. However, the current results demonstrate that variability in voice processing cortical regions in relation to HP exists without motor activity.

It is possible that theories proposing divergent involvement of speech production and perception mechanisms in AVH may be not mutually exclusive. Experiences of people who hallucinate are diverse. As theories of HP become more specific and concrete, they may become less well aligned with the phenomenology of the hallucinator. Therefore, hallucinatory experience might be best characterized by multiple subtypes, to which specific theories might apply better than others (Jones, 2010). For example, models describing the phenomenology of voice hearing ascribe the top-down contribution of intrusive memories and thoughts to the quality of false perception experiences (Hugdahl, 2015; Upthegrove et al., 2016; Bohlken et al., 2017; Ćurčić-Blake et al., 2017). A core abnormality in brain function central to the emergence of false perceptions likely rests in the interactive process of top-down predictions and bottom-up sensory input (Allen et al., 2008; Hugdahl, 2009, 2015; Kowalski et al., 2021). Regarding perceptual salience, bottom-up hypersensitivity to sensory input is congruent with established computation neuroscience accounts of predictive coding in false perceptions (Sterzer et al., 2018). Here, weighted top-down predictions and bottom-up explanations of sensation interact along a hierarchical network, constantly updating via Bayesian inference to form the most reliable percept (Friston, 2005, 2012; Fletcher and Frith, 2009; Feldman and Friston, 2010; Hohwy, 2017). When internal prediction signals are weighted too strongly, one "senses what they expect." Moreover, when the top-down input is too strong, the threshold for active perception may be reached under minimal sensory input. However, the self-monitoring theory posits a delayed or absent prediction signal resulting in increased activation of sensory cortical regions and is therefore in apparent conflict with the former account (Corlett et al., 2019; Leptourgos and Corlett, 2020). These expectations could operate on separate time scales, at different levels of the information processing hierarchy, or simply serve two different functions in hallucinations (Thakkar et al., 2021).

The role of perceptual salience in a multistage process leading to false perceptions has gathered substantial support in functional neuroimaging. Namely, research into large-scale functional brain networks has provided a resting-state hypothesis, outlining brain states serving as a predisposition for hallucinations, including voice hearing (Northoff and Qin, 2011; Northoff, 2014). While at rest, activation of the salience network, under conditions of irrelevant stimuli, may interrupt the Default Mode Network (DMN) and engage active sensory processing (Alderson-Day et al., 2015, 2016; Schmidt et al., 2015). The salience network therefore operates as a switch between the DMN and central executive network and how attention is directed toward incoming sensations, constituting a triple network model (TMN) subserving the advent of hallucinatory experience (Menon, 2011). Although we did not acquire behavioral data from the participants with ratings of perceived salience while listening to stimuli during scanning, we suggest that the change in brain activity that we observed in the right pSTG is indicative of the TMN in response to voice stimuli.

Hierarchical voice network processing

Voices are processed along a series of bilateral voice patches in the posterior, middle, and anterior STG. These temporal voice areas are reliably identified by a standardized TVA localizer task (Pernet et al., 2015). Participants with greater HP displayed increased right pSTG activation in response to vocal stimuli. Activity in this region may reflect sensitivity to low-level acoustic features during early stages of voice processing (Griffiths and Warren, 2004; Warren J. D. et al., 2005; Warren J. E. et al., 2005). Furthermore, the pSTG is not specialized for voice processing per se, and likely plays a broader role in extracting spectro-temporal acoustic features from complex sounds, of which voices are an example. However, activation in these regions preferentially responds to salient stimuli, such as voices, over and above other similarly complex environmental sounds (Pernet et al., 2007).

In terms of the salience hypothesis for hallucinatory experience, the assignment of salience to irrelevant, neutral, events must be considered in terms of the paralinguistic factors which may be involved. Indeed, the phenomenology of AVH is often marked by prominent paralinguistic features in the identity and emotional valence of the hallucinated speaker (Stephane et al., 2003; Larøi and Woodward, 2007; Badcock and Chhabra, 2013; McCarthy-Jones et al., 2014). Individuals who experience hallucinations often express difficulty in discerning the identity of veridical voices. For example, in schizophrenia patients who experience hallucinations, there is a bias to externalize voices to another person (Johns et al., 2001; Allen et al., 2007b; Mechelli et al., 2007; Pinheiro et al., 2016, 2017). Likewise, severity of AVH in patients is increasingly altered by emotional processing (Rossell and Boundy, 2005; Shea et al., 2007; Alba-Ferrara et al., 2013; Tseng et al., 2013). The role of salience may be influential in perceptions of speaker identity, as misattributions are more prevalent for emotional stimuli (Ditman and Kuperberg, 2005; Costafreda et al., 2008; Pinheiro et al., 2016, 2017). However, the effects of emotional valence in perceiving voice identity for people prone to false

perceptions of voices has not shown clear consensus (i.e., Brookwell et al., 2013). Comparisons of AVH severity in patients with schizophrenia with judgments of speaker identity have indicated an increasing proneness to externalize voices with negative content (Allen et al., 2004; Pinheiro et al., 2016). In non-clinical groups, the involvement of salient emotional features in voices is less clear. For example, higher levels of HP in the general population are not associated with atypical evaluation of emotional valence in words or vocalizations (Pinheiro et al., 2019). However, it has been indicated that non-clinical individuals prone to voice hearing require stronger emotional information to consider a stimulus as emotional (Amorim et al., 2021) or may allocate similar attention to voices irrespective of their emotional salience (Castiajo and Pinheiro, 2021). Future research is required into how variability in perceived salience of speaker-related features may affect processing in the hierarchical voice network and, in particular, how posterior STG activity related to HP may be influenced.

In addition to the TVA findings, the localizer task often provides a subset of extra-temporal regions indicating an extended voice processing network (Pernet et al., 2015). In our sample, extra-temporal peak activations were ascribed to bilateral inferior frontal and right hemisphere premotor cortex. Prefrontal involvement of the left IFC is commonly found in voice perception, with different subregions serving various functions. For example, the pars orbitalis is involved in processing semantic and emotional information (Belyk et al., 2017). Here, the left IFC peak was found in Broca's area, which has been theorized to represent mirror neuron activity which may be useful in guiding conversational turn-taking (Rizzolatti and Craighero, 2004; Grafton and Hamilton, 2007; Kilner et al., 2007). Likewise, precentral motor regions are involved in the perception and production of speech (Wilson et al., 2004; Pulvermüller et al., 2006; Cheung et al., 2016). This could explain speech production region activity sometimes reported during AVH (Jardri et al., 2011; Kühn and Gallinat, 2012; Zmigrod et al., 2016). However, self-monitoring theories take this as evidence for top-down inner speech signals guiding the perceived hallucinatory voice. Notably, transcranial direct-current stimulation targeting a fronto-parietal sensorimotor network is an effective treatment for the alleviation of AVH in patients with schizophrenia (Yang et al., 2019). In our *post hoc* analysis, the right IFC ROI shows an intriguing negative correlation to HP, however, only for non-voice sounds. The right IFC may serve a role in salience processing, for example in recognizing salient cues in voice signals (Johnstone et al., 2006; Bestelmeyer et al., 2012; Charest et al., 2013; Johns et al., 2015; Johnson et al., 2021). Additionally, this area shares a high functional integration with temporal regions serving voice perception and may assist successful voice recognition (Aglieri et al., 2018). Although this finding is difficult to interpret on its own, it may indicate a decrease in salience attribution for environmental sounds during a voice perception task. This

may indicate not only an HP-related salience bias affecting the sensitivity of cortical responses to voice sounds, but also a general bias away from non-voice sounds between hypersalient responses to intermittent voice stimuli.

Limitations and recommendations

We identify a number of limitations within the current study and provide suggestions for future research. First, although the use of the established TVA localizer task facilitated the testing of our hypotheses regarding an early hypersensitivity to voice sounds, it did not preclude further investigation into how more complex stages of the voice processing hierarchy may relate to HP. Specifically, BOLD responses from this task are averaged across the trials containing different types of voice stimuli. This implies that signals extracted from ROIs serving different functional roles in voice processing, e.g., emotion or identity, do not represent the processing of specific features, but rather constitute a generalized voice detection signal. Second, in this study, behavioral measures of perceived stimulus salience were not collected. Therefore, interpretations of a salience bias attributed to increased functional brain responses cannot be directly linked to the subjective perception of the participants. Third, participants in the current study were sampled from a relatively homogenous sample of university students, similar in age, ethnicity, and cultural backgrounds. Due to the uneven distribution of environmental risk factors for psychotic symptoms throughout the population (Johns and van Os, 2001, DeRosse and Karlsgodt, 2015; Baumeister et al., 2017), our sample may unintentionally capture a set of protective factors. To address these limitations in future studies, we suggest a two-step procedure using a novel task that systematically varies paralinguistic voice features. This may allow investigations into how hierarchical processing downstream of initial HP-related hypersensitivity may influence responses to the perceived emotion or identity of the speaker. Furthermore, behavioral appraisals of perceived salience may be included to compare fMRI response patterns and HP scores. Finally, subsequent research may benefit from an increased sample size and diversity, including a structured collection of additional demographic data and associated environmental risk factors as possible covariates for HP-related brain changes.

Conclusion

We observed that HP is positively correlated with increased activation in the right pSTG in response to passively heard voices. This suggests a hypersensitivity associated with a propensity to hallucinate in a region of the brain which extracts low-level acoustic features from complex auditory signals. The right pSTG comprises the early processing of

voice signals along the paralinguistic information pathway of the cortical voice processing network. We propose that this increases activity in response to voices represents a perceptual salience bias as a precursor for the emergence of hallucinations. This interpretation is in line with functional network models that posit abnormal engagement of a salience network during irrelevant stimulus exposure as the underlying neurocognitive mechanism of false perceptions. Furthermore, the current findings conflict with self-monitoring accounts of inner speech models that propose a critical role of voice production regions in the inception of AVH. We have demonstrated that HP is associated with right pSTG activation driven by external auditory signals. Although we do not reject self-monitoring accounts, we suggest that a state of cortical hypersensitivity to irrelevant sensory input may be the first step in the emergence of a hallucinatory experience, possibly followed by the influence of top-down signals such as inner speech, memory, and thought that together contribute to the phenomenology of AVH.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethical Review Committee of the Faculty of Psychology and Neuroscience at Maastricht University. The patients/participants provided their written informed consent to participate in this study.

Author contributions

JJ conceptualized and carried out experiment, performed the analyses, and wrote manuscript with input from all authors. MB verified the analytical methods. SK, AP, MS, and MB conceptualized and interpreted the results. SK and AP provided the original idea for project and secured funding. All authors contributed to the article and approved the submitted version.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnhum.2022.859731/full#supplementary-material

References

Aglieri, V., Chaminade, T., Takerkart, S., and Belin, P. (2018). Functional connectivity within the voice perception network and its behavioural relevance. *Neuroimage* 183, 356–365. doi: 10.1016/j.neuroimage.2018.08.011

Alba-Ferrara, L., De Erausquin, G. A., Hirnstein, M., Weis, S., and Hausmann, M. (2013). Emotional prosody modulates attention in schizophrenia patients with hallucinations. *Front. Hum. Neurosci.* 7:59. doi: 10.3389/fnhum.2013.00059

Alderson-Day, B., Diederen, K., Fernyhough, C., Ford, J. M., Horga, G., Margulies, D. S., et al. (2016). Auditory hallucinations and the brain's resting-state networks: findings and methodological observations. *Schizophr. Bull.* 42, 1110–1123. doi: 10.1093/schbul/sbw078

Alderson-Day, B., McCarthy-Jones, S., and Fernyhough, C. (2015). Hearing voices in the resting brain: a review of intrinsic functional connectivity research on auditory verbal hallucinations. *Neurosci. Biobehav. Rev.* 55, 78–87. doi: 10.1016/j.neubiorev.2015.04.016

Allen, P. P., Johns, L. C., Fu, C. H., Broome, M. R., Vythelingum, G. N., and McGuire, P. K. (2004). Misattribution of external speech in patients with hallucinations and delusions. *Schizophr. Res.* 69, 277–287. doi: 10.1016/j.schres.2003.09.008

Allen, P., Aleman, A., and Mcguire, P. K. (2007a). Inner speech models of auditory verbal hallucinations: evidence from behavioural and neuroimaging studies. *Int. Rev. Psychiatry* 19, 407–415. doi: 10.1080/09540260701486498

Allen, P., Amaro, E., Fu, C. H., Williams, S. C., Brammer, M. J., Johns, L. C., et al. (2007b). Neural correlates of the misattribution of speech in schizophrenia. *Br. J. Psychiatry* 190, 162–169. doi: 10.1192/bjp.bp.106.025700

Allen, P., Larøi, F., McGuire, P. K., and Aleman, A. (2008). The hallucinating brain: a review of structural and functional neuroimaging studies of hallucinations. *Neurosci. Biobehav. Rev.* 32, 175–191. doi: 10.1016/j.neubiorev.2007.07.012

Amorim, M., Roberto, M. S., Kotz, S. A., and Pinheiro, A. P. (2021). The perceived salience of vocal emotions is dampened in non-clinical auditory verbal hallucinations. *Cogn. Neuropsychiatry* 27, 169–182. doi: 10.1080/13546805.2021.1949972

Anthony, D. (2004). The cognitive neuropsychiatry of auditory verbal hallucinations: an overview. *Cogn. Neuropsychiatry* 9, 107–123. doi: 10.1080/13546800344000183

Badcock, J. C., and Chhabra, S. (2013). Voices to reckon with: perceptions of voice identity in clinical and non-clinical voice hearers. *Front. Hum. Neurosci.* 7:114. doi: 10.3389/fnhum.2013.00114

Bauer, S. M., Schanda, H., Karakula, H., Olajossy-Hilkesberger, L., Rudaleviciene, P., Okribelashvili, N., et al. (2011). Culture and the prevalence of hallucinations in schizophrenia. *Compr. Psychiatry* 52, 319–325. doi: 10.1016/j.comppsych.2010.06.008

Baumeister, D., Sedgwick, O., Howes, O., and Peters, E. (2017). Auditory verbal hallucinations and continuum models of psychosis: a systematic review of the healthy voice-hearer literature. *Clin. Psychol. Rev.* 51, 125–141. doi: 10.1016/j.cpr.2016.10.010

Beavan, V., Read, J., and Cartwright, C. (2011). The prevalence of voice-hearers in the general population: a literature review. *J. Ment. Health* 20, 281–292. doi: 10.3109/09638237.2011.562262

Belin, P., and Zatorre, R. J. (2003). Adaptation to speaker's voice in right anterior temporal lobe. *Neuroreport* 14, 2105–2109. doi: 10.1097/00001756-200311140-00019

Belin, P., Fecteau, S., and Bedard, C. (2004). Thinking the voice: neural correlates of voice perception. *Trends Cogn. Sci.* 8, 129–135. doi: 10.1016/j.tics.2004.01.008

Belin, P., Zatorre, R. J., Hoge, R., Evans, A. C., and Pike, B. (1999). Event-related fMRI of the auditory cortex. *Neuroimage* 10, 417–429. doi: 10.1006/nimg.1999.0480

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., and Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature* 403, 309–312. doi: 10.1038/35002078

Belyk, M., Brown, S., Lim, J., and Kotz, S. A. (2017). Convergence of semantics and emotional expression within the IFG pars orbitalis. *Neuroimage* 156, 240–248. doi: 10.1016/j.neuroimage.2017.04.020

Bentall, R. P. (1990). The illusion of reality: a review and integration of psychological research on hallucinations. *Psychol. Bull.* 107:82. doi: 10.1037/0033-2909.107.1.82

Bestelmeyer, P. E., Latinus, M., Bruckert, L., Rouger, J., Crabbe, F., and Belin, P. (2012). Implicitly perceived vocal attractiveness modulates prefrontal cortex activity. *Cereb. Cortex* 22, 1263–1270. doi: 10.1093/cercor/bhr204

Bohlken, M. M., Hugdahl, K., and Sommer, I. E. (2017). Auditory verbal hallucinations: neuroimaging and treatment. *Psychol. Med.* 47, 199–208. doi: 10.1017/S003329171600115X

Brett, M., Anton, J., Valabregue, R., and Poline, J. (2002). Region of interest analysis using an SPM toolbox [abstract]. *Paper Presented at the 8th International Conference on Functional Mapping of the Human Brain*, Sendai.

Brookwell, M. L., Bentall, R. P., and Varese, F. (2013). Externalizing biases and hallucinations in source-monitoring, self-monitoring and signal detection studies: a meta-analytic review. *Psychol. Med.* 43, 2465–2475. doi: 10.1017/S0033291712002760

Castiajo, P., and Pinheiro, A. P. (2021). Acoustic salience in emotional voice perception and its relationship with hallucination proneness. *Cogn. Affect. Behav. Neurosci.* 21, 412–425. doi: 10.3758/s13415-021-00864-2

Charest, I., Pernet, C., Latinus, M., Crabbe, F., and Belin, P. (2013). Cerebral processing of voice gender studied using a continuous carryover fMRI design. *Cereb. Cortex* 23, 958–966. doi: 10.1093/cercor/bhs090

Cheung, C., Hamilton, L. S., Johnson, K., and Chang, E. F. (2016). The auditory representation of speech sounds in human motor cortex. *Elife* 5:e12577. doi: 10.7554/eLife.12577

Copolov, D. L., Seal, M. L., Maruff, P., Ulusoy, R., Wong, M. T., Tochon-Danguy, H. J., et al. (2003). Cortical activation associated with the experience of auditory hallucinations and perception of human speech in schizophrenia: a PET correlation study. *Psychiatry Res. Neuroimaging* 122, 139–152. doi: 10.1016/S0925-4927(02)00121-X

Corlett, P. R., Horga, G., Fletcher, P. C., Alderson-Day, B., Schmack, K., and Powers, A. R. III (2019). Hallucinations and strong priors. *Trends Cogn. Sci.* 23, 114–127. doi: 10.1016/j.tics.2018.12.001

Costafreda, S. G., Brébion, G., Allen, P., McGuire, P. K., and Fu, C. H. (2008). Affective modulation of external misattribution bias in source monitoring in schizophrenia. *Psychol. Med.* 38, 821–824. doi: 10.1017/S0033291708003243

Ćurčić-Blake, B., Ford, J. M., Hubl, D., Orlov, N. D., Sommer, I. E., Waters, F., et al. (2017). Interaction of language, auditory and memory brain networks in auditory verbal hallucinations. *Prog. Neurobiol.* 148, 1–20. doi: 10.1016/j.pneurobio.2016.11.002

Daalman, K., Boks, M. P., Diederen, K. M., de Weijer, A. D., Blom, J. D., Kahn, R. S., et al. (2011). The same or different? A phenomenological comparison of auditory verbal hallucinations in healthy and psychotic individuals. *J. Clin. Psychiatry* 72, 320–325. doi: 10.4088/JCP.09m05797yel

Daalman, K., Verkooijen, S., Derks, E. M., Aleman, A., and Sommer, I. E. C. (2012). The influence of semantic top-down processing in auditory verbal hallucinations. *Schizophr. Res.* 139, 82–86. doi: 10.1016/j.schres.2012.06.005

de Boer, J. N., Heringa, S. M., van Dellen, E., Wijnen, F. N. K., and Sommer, I. E. C. (2016). A linguistic comparison between auditory verbal hallucinations in patients with a psychotic disorder and in nonpsychotic individuals: not just what the voices say, but how they say it. *Brain Lang.* 162, 10–18. doi: 10.1016/j.bandl.2016.07.011

de Leede-Smith, S., and Barkus, E. (2013). A comprehensive review of auditory verbal hallucinations: lifetime prevalence, correlates and mechanisms in healthy and clinical individuals. *Front. Hum. Neurosci.* 7:367. doi: 10.3389/fnhum.2013.00367

Deneve, S., and Jardri, R. (2016). Circular inference: mistaken belief, misplaced trust. *Curr. Opin. Behav. Sci.* 11, 40–48. doi: 10.1016/j.cobeha.2016.04.001

DeRosse, P., and Karlsgodt, K. H. (2015). Examining the psychosis continuum. *Curr. Behav. Neurosci. Rep.* 2, 80–89. doi: 10.1007/s40473-015-0040-7

Ditman, T., and Kuperberg, G. R. (2005). A source-monitoring account of auditory verbal hallucinations in patients with schizophrenia. *Harv. Rev. Psychiatry* 13, 280–299. doi: 10.1080/10673220500326391

Escartí, M. J., de la Iglesia-Vayá, M., Martí-Bonmatí, L., Robles, M., Carbonell, J., Lull, J. J., et al. (2010). Increased amygdala and parahippocampal gyrus activation in schizophrenic patients with auditory hallucinations: an fMRI study using independent component analysis. *Schizophr. Res.* 117, 31–41. doi: 10.1016/j.schres.2009.12.028

Feinberg, I. (1978). Efference copy and corollary discharge: implications for thinking and its disorders. *Schizophr. Bull.* 4:636. doi: 10.1093/schbul/4.4.636

Feldman, H., and Friston, K. (2010). Attention, uncertainty, and free-energy. *Front. Hum. Neurosci.* 4:215. doi: 10.3389/fnhum.2010.00215

Fletcher, P. C., and Frith, C. D. (2009). Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat. Rev. Neurosci.* 10, 48–58. doi: 10.1038/nrn2536

Fonov, V. S., Evans, A. C., McKinstry, R. C., Almli, C. R., and Collins, D. L. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *Neuroimage* 47:S102. doi: 10.1016/S1053-8119(09)70884-5

Ford, J. M., and Mathalon, D. H. (2019). Efference copy, corollary discharge, predictive coding, and psychosis. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 4, 764–767. doi: 10.1016/j.bpsc.2019.07.005

Formisano, E., De Martino, F., Bonte, M., and Goebel, R. (2008). " Who" is saying" what"? Brain-based decoding of human voice and speech. *Science* 322, 970–973. doi: 10.1126/science.1164318

Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R. Soc. B Biol. Sci.* 360, 815–836.

Friston, K. (2012). Prediction, perception and agency. *Int. J. Psychophysiol.* 83, 248–252. doi: 10.1016/j.ijpsycho.2011.11.014

Frith, C. D., and Done, D. J. (1988). Towards a neuropsychology of schizophrenia. *Br. J. Psychiatry* 153, 437–443. doi: 10.1192/bjp.153.4.437

Grafton, S. T., and Hamilton, A. F. D. C. (2007). Evidence for a distributed hierarchy of action representation in the brain. *Hum. Mov. Sci.* 26, 590–616. doi: 10.1016/j.humov.2007.05.009

Gregory, D. (2016). Inner speech, imagined speech, and auditory verbal hallucinations. *Rev. Philos. Psychol.* 7, 653–673. doi: 10.1007/s13164-015-0274-z

Griffiths, T. D., and Warren, J. D. (2004). What is an auditory object? *Nat. Rev. Neurosci.* 5, 887–892. doi: 10.1038/nrn1538

Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, A. Q., Elliott, M. R., et al. (1999). "Sparse" temporal sampling in auditory fMRI. *Hum. Brain Mapp.* 7, 213–223. doi: 10.1002/(SICI)1097-0193199973<213::AID-HBM5>3.0.CO;2-N

Hohwy, J. (2017). Priors in perception: top-down modulation, Bayesian perceptual learning rate, and prediction error minimization. *Conscious. Cogn.* 47, 75–85. doi: 10.1016/j.concog.2016.09.004

Hugdahl, K. (2009). "Hearing voices": auditory hallucinations as failure of top-down control of bottom-up perceptual processes. *Scand. J. Psychol.* 50, 553–560. doi: 10.1111/j.1467-9450.2009.00775.x

Hugdahl, K. (2015). Auditory hallucinations: a review of the ERC "VOICE" project. *World J. Psychiatry* 5:193. doi: 10.5498/wjp.v5.i2.193

Hugdahl, K., and Sommer, I. E. (2018). Auditory verbal hallucinations in schizophrenia from a levels of explanation perspective. *Schizophr. Bull.* 44, 234–241. doi: 10.1093/schbul/sbx142

Jardri, R., Hugdahl, K., Hughes, M., Brunelin, J., Waters, F., Alderson-Day, B., et al. (2016). Are hallucinations due to an imbalance between excitatory and inhibitory influences on the brain? *Schizophr. Bull.* 42, 1124–1134. doi: 10.1093/schbul/sbw075

Jardri, R., Pouchet, A., Pins, D., and Thomas, P. (2011). Cortical activations during auditory verbal hallucinations in schizophrenia: a coordinate-based meta-analysis. *Am. J. Psychiatry* 168, 73–81. doi: 10.1176/appi.ajp.2010.09101522

Johns, L. C. (2005). Hallucinations in the general population. *Curr. Psychiatry Rep.* 7, 162–167. doi: 10.1006/nimg.2002.1132

Johns, L. C., and van Os, J. (2001). The continuity of psychotic experiences in the general population. *Clin. Psychol. Rev.* 21, 1125–1141. doi: 10.1016/S0272-7358(01)00103-9

Johns, A. B., Farrall, A. J., Belin, P., and Pernet, C. R. (2015). Hemispheric association and dissociation of voice and speech information processing in stroke. *Cortex* 71, 232–239. doi: 10.1016/j.cortex.2015.07.004

Johns, L. C., Kompus, K., Connell, M., Humpston, C., Lincoln, T. M., Longden, E., et al. (2014). Auditory verbal hallucinations in persons with and without a need for care. *Schizophr. Bull.* 40(Suppl. 4), S255–S264. doi: 10.1093/schbul/sbu005

Johns, L. C., Rossell, S., Frith, C., Ahmad, F., Hemsley, D., Kuipers, E., et al. (2001). Verbal self-monitoring and auditory verbal hallucinations in patients with schizophrenia. *Psychol. Med.* 31, 705–715. doi: 10.1017/S0033291701003774

Jones, S. R. (2010). Do we need multiple models of auditory verbal hallucinations? Examining the phenomenological fit of cognitive and neurological models. *Schizophr. Bull.* 36, 566–575. doi: 10.1093/schbul/sbn129

Jones, S. R., and Fernyhough, C. (2007a). Thought as action: inner speech, self-monitoring, and auditory verbal hallucinations. *Conscious. Cogn.* 16, 391–399. doi: 10.1016/j.concog.2005.12.003

Jones, S. R., and Fernyhough, C. (2007b). Neural correlates of inner speech and auditory verbal hallucinations: a critical review and theoretical integration. *Clin. Psychol. Rev.* 27, 140–154. doi: 10.1016/j.cpr.2006.10.001

Johnson, J. F., Belyk, M., Schwartze, M., Pinheiro, A. P., and Kotz, S. A. (2021). Expectancy changes the self-monitoring of voice identity. *Eur. J. Neurosci.* 53, 2681–2695. doi: 10.1111/ejn.15162

Johnstone, T., Van Reekum, C. M., Oakes, T. R., and Davidson, R. J. (2006). The voice of emotion: an FMRI study of neural responses to angry and happy vocal expressions. *Soc. Cogn. Affect. Neurosci.* 1, 242–249. doi: 10.1093/scan/nsl027

Kelleher, I., and Cannon, M. (2011). Psychotic-like experiences in the general population: characterizing a high-risk group for psychosis. *Psychol. Med.* 41, 1–6. doi: 10.1017/S0033291710001005

Kelleher, I., Jenner, J. A., and Cannon, M. (2010). Psychotic symptoms in the general population–an evolutionary perspective. *Br. J. Psychiatry* 197, 167–169. doi: 10.1192/bjp.bp.109.076018

Kilner, J. M., Friston, K. J., and Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cogn. Process.* 8, 159–166. doi: 10.1007/s10339-007-0170-2

Kompus, K., Løberg, E. M., Posserud, M. B., and Lundervold, A. J. (2015). Prevalence of auditory hallucinations in Norwegian adolescents: results from a population-based study. *Scand. J. Psychol.* 56, 391–396. doi: 10.1111/sjop.12219

Kompus, K., Westerhausen, R., and Hugdahl, K. (2011). The "paradoxical" engagement of the primary auditory cortex in patients with auditory verbal hallucinations: a meta-analysis of functional neuroimaging studies. *Neuropsychologia* 49, 3361–3369. doi: 10.1016/j.neuropsychologia.2011.08.010

Kowalski, J., Aleksandrowicz, A., Dąbkowska, M., and Gawęda, Ł (2021). Neural correlates of aberrant salience and source monitoring in schizophrenia and at-risk mental states—a systematic review of fMRI studies. *J. Clin. Med.* 10:4126. doi: 10.3390/jcm10184126

Kühn, S., and Gallinat, J. (2012). Quantitative meta-analysis on state and trait aspects of auditory verbal hallucinations in schizophrenia. *Schizophr. Bull.* 38, 779–786. doi: 10.1093/schbul/sbq152

Larøi, F., and Van Der Linden, M. (2005). Nonclinical participants' reports of hallucinatory experiences. *Can. J. Behav. Sci.* 37:33. doi: 10.1037/h0087243

Larøi, F., and Woodward, T. S. (2007). Hallucinations from a cognitive perspective. *Harv. Rev. Psychiatry* 15, 109–117. doi: 10.1080/10673220701401993

Larøi, F., Marczewski, P., and Van der Linden, M. (2004). Further evidence of the multi-dimensionality of hallucinatory predisposition: factor structure of a modified version of the Launay-Slade hallucinations scale in a normal sample. *Eur. Psychiatry* 19, 15–20. doi: 10.1016/S0924-9338(03)00028-2

Larøi, F., Sommer, I. E., Blom, J. D., Fernyhough, C., Ffytche, D. H., Hugdahl, K., et al. (2012). The characteristic features of auditory verbal hallucinations in clinical and nonclinical groups: state-of-the-art overview and future directions. *Schizophr. Bull.* 38, 724–733. doi: 10.1093/schbul/sbs061

Lavan, N., Burton, A. M., Scott, S. K., and McGettigan, C. (2019). Flexible voices: identity perception from variable vocal signals. *Psychonom. Bull. Rev.* 26, 90–102. doi: 10.3758/s13423-018-1497-7

Leptourgos, P., and Corlett, P. R. (2020). Embodied predictions, agency, and psychosis. *Front. Big Data* 3:27. doi: 10.3389/fdata.2020.00027

Leptourgos, P., Denève, S., and Jardri, R. (2017). Can circular inference relate the neuropathological and behavioral aspects of schizophrenia? *Curr. Opin. Neurobiol.* 46, 154–161. doi: 10.1016/j.conb.2017.08.012

Linscott, R. J., and Van Os, J. (2013). An updated and conservative systematic review and meta-analysis of epidemiological evidence on psychotic experiences in children and adults: on the pathway from proneness to persistence to dimensional expression across mental disorders. *Psychol. Med.* 43:1133. doi: 10.1017/S0033291712001626

Martí-Bonmatí, L., Lull, J. J., García-Martí, G., Aguilar, E. J., Moratal-Pérez, D., Poyatos, C., et al. (2007). Chronic auditory hallucinations in schizophrenic patients: MR analysis of the coincidence between functional and morphologic abnormalities. *Radiology* 244, 549–556. doi: 10.1148/radiol.2442060727

McCarthy, P. (2022). *FSLeyes (1.4.0)*. Zenodo. doi: 10.5281/zenodo.6511596

McCarthy-Jones, S., Trauer, T., Mackinnon, A., Sims, E., Thomas, N., and Copolov, D. L. (2014). A new phenomenological survey of auditory hallucinations: evidence for subtypes and implications for theory and practice. *Schizophr. Bull.* 40, 231–235. doi: 10.1093/schbul/sbs156

McGrath, J. J., Saha, S., Al-Hamzawi, A., Alonso, J., Bromet, E. J., Bruffaerts, R., et al. (2015). Psychotic experiences in the general population: a cross-national analysis based on 31 261 respondents from 18 countries. *JAMA Psychiatry* 72, 697–705. doi: 10.1001/jamapsychiatry.2015.0575

Mechelli, A., Allen, P., Amaro, E. Jr., Fu, C. H., Williams, S. C., Brammer, M. J., et al. (2007). Misattribution of speech and impaired connectivity in patients with auditory verbal hallucinations. *Hum. Brain Mapp.* 28, 1213–1222. doi: 10.1002/hbm.20341

Menon, V. (2011). Large-scale brain networks and psychopathology: a unifying triple network model. *Trends Cogn. Sci.* 15, 483–506. doi: 10.1016/j.tics.2011.08.003

Menon, V., and Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Struct. Funct.* 214, 655–667. doi: 10.1007/s00429-010-0262-0

Miyata, J. (2019). Toward integrated understanding of salience in psychosis. *Neurobiol. Dis.* 131:104414. doi: 10.1016/j.nbd.2019.03.002

Nakamura, K., Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., et al. (2001). Neural substrates for recognition of familiar voices: a PET study. *Neuropsychologia* 39, 1047–1054. doi: 10.1016/S0028-3932(01)00037-9

Northoff, G. (2014). Are auditory hallucinations related to the brain's resting state activity? A 'neurophenomenal resting state hypothesis'. *Clin. Psychopharmacol. Neurosci.* 12:189. doi: 10.9758/cpn.2014.12.3.189

Northoff, G., and Qin, P. (2011). How can the brain's resting state activity generate hallucinations? A 'resting state hypothesis' of auditory verbal hallucinations. *Schizophr. Res.* 127, 202–214. doi: 10.1016/j.schres.2010.11.009

Palaniyappan, L., and Liddle, P. F. (2012). Does the salience network play a cardinal role in psychosis? An emerging hypothesis of insular dysfunction. *J. Psychiatry Neurosci.* 37, 17–27. doi: 10.1503/jpn.100176

Parellada, E., Lomena, F., Font, M., Pareto, D., Gutierrez, F., Simo, M., et al. (2008). Fluordeoxyglucose-PET study in first-episode schizophrenic patients during the hallucinatory state, after remission and during linguistic–auditory activation. *Nuclear Med. Commun.* 29, 894–900. doi: 10.1097/MNM.0b013e328302cd10

Pernet, C., Schyns, P. G., and Demonet, J. F. (2007). Specific, selective or preferential: comments on category specificity in neuroimaging. *Neuroimage* 35, 991–997. doi: 10.1016/j.neuroimage.2007.01.017

Pernet, C. R., McAleer, P., Latinus, M., Gorgolewski, K. J., Charest, I., Bestelmeyer, P. E., et al. (2015). The human voice areas: spatial organization and inter-individual variability in temporal and extra-temporal cortices. *Neuroimage* 119, 164–174. doi: 10.1016/j.neuroimage.2015.06.050

Pinheiro, A. P., Farinha-Fernandes, A., Roberto, M. S., and Kotz, S. A. (2019). Self-voice perception and its relationship with hallucination predisposition. *Cogn. Neuropsychiatry* 24, 237–255. doi: 10.1080/13546805.2019.1621159

Pinheiro, A. P., Rezaii, N., Rauber, A., and Niznikiewicz, M. (2016). Is this my voice or yours? The role of emotion and acoustic quality in self-other voice discrimination in schizophrenia. *Cogn. Neuropsychiatry* 21, 335–353. doi: 10.1080/13546805.2016.1208611

Pinheiro, A. P., Rezaii, N., Rauber, A., Nestor, P. G., Spencer, K. M., and Niznikiewicz, M. (2017). Emotional self–other voice processing in schizophrenia and its relationship with hallucinations: ERP evidence. *Psychophysiology* 54, 1252–1265. doi: 10.1111/psyp.12880

Pulvermüller, F., Huss, M., Kherif, F., Moscoso del Prado Martin, F., Hauk, O., and Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proc. Natl. Acad. Sci. U.S.A.* 103, 7865–7870. doi: 10.1073/pnas.0509989103

Reininghaus, U., Kempton, M. J., Valmaggia, L., Craig, T. K., Garety, P., Onyejiaka, A., et al. (2016). Stress sensitivity, aberrant salience, and threat anticipation in early psychosis: an experience sampling study. *Schizophr. Bull.* 42, 712–722. doi: 10.1093/schbul/sbv190

Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192. doi: 10.1146/annurev.neuro.27.070203.144230

Rollins, C. P., Garrison, J. R., Simons, J. S., Rowe, J. B., O'Callaghan, C., Murray, G. K., et al. (2019). Meta-analytic evidence for the plurality of mechanisms in transdiagnostic structural MRI studies of hallucination status. *EClinicalMedicine* 8, 57–71. doi: 10.1016/j.eclinm.2019.01.012

Rossell, S. L., and Boundy, C. L. (2005). Are auditory–verbal hallucinations associated with auditory affective processing deficits? *Schizophr. Res.* 78, 95–106. doi: 10.1016/j.schres.2005.06.002

Rousseeuw, P. J., and Hubert, M. (2011). Robust statistics for outlier detection. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* 1, 73–79. doi: 10.1002/widm.2

Schirmer, A., and Kotz, S. A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends Cogn. Sci.* 10, 24–30. doi: 10.1016/j.tics.2005.11.009

Schmidt, A., Diwadkar, V. A., Smieskova, R., Harrisberger, F., Lang, U. E., McGuire, P., et al. (2015). Approaching a network connectivity-driven classification of the psychosis continuum: a selective review and suggestions for future research. *Front. Hum. Neurosci.* 8:1047. doi: 10.3389/fnhum.2014.01047

Shea, T. L., Sergejew, A. A., Burnham, D., Jones, C., Rossell, S. L., Copolov, D. L., et al. (2007). Emotional prosodic processing in auditory hallucinations. *Schizophr. Res.* 90, 214–220. doi: 10.1016/j.schres.2006.09.021

Simons, C. J., Tracy, D. K., Sanghera, K. K., O'Daly, O., Gilleen, J., Krabbendam, L., et al. (2010). Functional magnetic resonance imaging of inner speech in schizophrenia. *Biol. Psychiatry* 67, 232–237. doi: 10.1016/j.biopsych.2009.09.007

Stephane, M., Thuras, P., Nasrallah, H., and Georgopoulos, A. P. (2003). The internal structure of the phenomenology of auditory verbal hallucinations. *Schizophr. Res.* 61, 185–193. doi: 10.1016/S0920-9964(03)00013-6

Sterzer, P., Adams, R. A., Fletcher, P., Frith, C., Lawrie, S. M., Muckli, L., et al. (2018). The predictive coding account of psychosis. *Biol. Psychiatry* 84, 634–643. doi: 10.1016/j.biopsych.2018.05.015

Swiney, L., and Sousa, P. (2014). A new comparator account of auditory verbal hallucinations: how motor prediction can plausibly contribute to the sense of agency for inner speech. *Front. Hum. Neurosci.* 8:675. doi: 10.3389/fnhum.2014.00675

Thakkar, K. N., Mathalon, D. H., and Ford, J. M. (2021). Reconciling competing mechanisms posited to underlie auditory verbal hallucinations. *Philos. Trans. R. Soc. B* 376:20190702. doi: 10.1098/rstb.2019.0702

Tracy, D. K., and Shergill, S. S. (2006). Imaging auditory hallucinations in schizophrenia. *Acta Neuropsychiatr.* 18, 71–78. doi: 10.1111/j.1601-5215.2006.00129.x

Tseng, H. H., Chen, S. H., Liu, C. M., Howes, O., Huang, Y. L., Hsieh, M. H., et al. (2013). Facial and prosodic emotion recognition deficits associate with specific clusters of psychotic symptoms in schizophrenia. *PLoS One* 8:e66571. doi: 10.1371/journal.pone.0066571

Uddin, L. Q. (2015). Salience processing and insular cortical function and dysfunction. *Nat. Rev. Neurosci.* 16, 55–61. doi: 10.1038/nrn3857

Upthegrove, R., Broome, M. R., Caldwell, K., Ives, J., Oyebode, F., and Wood, S. J. (2016). Understanding auditory verbal hallucinations: a systematic review of current evidence. *Acta Psychiatr. Scand.* 133, 352–367. doi: 10.1111/acps.12531

van Lutterveld, R., Diederen, K. M., Koops, S., Begemann, M. J., and Sommer, I. E. (2013). The influence of stimulus detection on activation patterns during auditory hallucinations. *Schizophr. Res.* 145, 27–32. doi: 10.1016/j.schres.2013.01.004

Van Os, J., Hanssen, M., Bijl, R. V., and Ravelli, A. (2000). Strauss (1969) revisited: a psychosis continuum in the general population? *Schizophr. Res.* 45, 11–20. doi: 10.1016/S0920-9964(99)00224-8

Verdoux, H., and van Os, J. (2002). Psychotic symptoms in non-clinical populations and the continuum of psychosis. *Schizophr. Res.* 54, 59–65. doi: 10.1016/S0920-9964(01)00352-8

von Kriegstein, K., and Giraud, A. L. (2004). Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *Neuroimage* 22, 948–955. doi: 10.1016/j.neuroimage.2004.02.020

von Kriegstein, K., Eger, E., Kleinschmidt, A., and Giraud, A. L. (2003). Modulation of neural responses to speech by directing attention to voices or verbal content. *Cogn. Brain Res.* 17, 48–55. doi: 10.1016/S0926-6410(03)00079-X

Warren, J. D., Jennings, A. R., and Griffiths, T. D. (2005). Analysis of the spectral envelope of sounds by the human brain. *Neuroimage* 24, 1052–1057. doi: 10.1016/j.neuroimage.2004.10.031

Warren, J. E., Wise, R. J., and Warren, J. D. (2005). Sounds do-able: auditory–motor transformations and the posterior temporal plane. *Trends Neurosci.* 28, 636–643. doi: 10.1016/j.tins.2005.09.010

Waters, F., and Fernyhough, C. (2017). Hallucinations: a systematic review of points of similarity and difference across diagnostic classes. *Schizophr. Bull.* 43, 32–43. doi: 10.1093/schbul/sbw132

Waters, F., Allen, P., Aleman, A., Fernyhough, C., Woodward, T. S., Badcock, J. C., et al. (2012). Auditory hallucinations in schizophrenia and nonschizophrenia populations: a review and integrated model of cognitive mechanisms. *Schizophr. Bull.* 38, 683–693. doi: 10.1093/schbul/sbs045

Weiss, A. P., and Heckers, S. (1999). Neuroimaging of hallucinations: a review of the literature. *Psychiatry Res. Neuroimaging* 92, 61–74. doi: 10.1016/S0925-4927(99)00041-4

Wilson, S. M., Saygin, A. P., Sereno, M. I., and Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nat. Neurosci.* 7, 701–702. doi: 10.1038/nn1263

Woodruff, P. W., Wright, I. C., Bullmore, E. T., Brammer, M., Howard, R. J., Williams, S. C., et al. (1997). Auditory hallucinations and the temporal cortical response to speech in schizophrenia: a functional magnetic resonance imaging study. *Am. J. Psychiatry* 154, 1676–1682. doi: 10.1176/ajp.154.12.1676

Yang, F., Fang, X., Tang, W., Hui, L., Chen, Y., Zhang, C., et al. (2019). Effects and potential mechanisms of transcranial direct current stimulation (tDCS) on auditory hallucinations: a meta-analysis. *Psychiatry Res.* 273, 343–349. doi: 10.1016/j.psychres.2019.01.059

Zhang, Y., Ding, Y., Huang, J., Zhou, W., Ling, Z., Hong, B., et al. (2021). Hierarchical cortical networks of "voice patches" for processing voices in human brain. *Proc. Natl. Acad. Sci. U.S.A.* 118, e2113887118. doi: 10.1073/pnas.2113887118

Zhuo, C., Jiang, D., Liu, C., Lin, X., Li, J., Chen, G., et al. (2019). Understanding auditory verbal hallucinations in healthy individuals and individuals with psychiatric disorders. *Psychiatry Res.* 274, 213–219. doi: 10.1016/j.psychres.2019.02.040

Zmigrod, L., Garrison, J. R., Carr, J., and Simons, J. S. (2016). The neural mechanisms of hallucinations: a quantitative meta-analysis of neuroimaging studies. *Neurosci. Biobehav. Rev.* 69, 113–123. doi: 10.1016/j.neubiorev.2016.05.037