Title: Impaired speaking-induced suppression in schizophrenia predicts auditory verbal hallucinations

Authors: Authors: Songyuan Tan[1], Yingxin Jia[1], Namasvi Jariwala[1,4], Alvincé Pongos[2], Kurtis Brent[2], John Houde[2], Srikantan Nagarajan[3], Karuna Subramaniam*[1]

[1] Department of Psychiatry, University of California, San Francisco, CA

[2] Department of Otolaryngology, University of California, San Francisco, San Francisco, CA

[3] Department of Radiology and Biomedical Imaging, University of California, San Francisco, San Francisco, CA

[4] Department of Psychology, Palo Alto University, CA

***Corresponding Author:** Karuna Subramaniam

Abstract

A successful efference copy self-prediction suppresses auditory signals in the primary auditory cortex (A1) is necessary for speakers to successfully compare auditory feedback during speech production with auditory feedback during passive listening, this is called speaker-induced suppression (SIS). The top-rank positive symptom in schizophrenic (SZ) patients, auditory verbal hallucination, for instance, is hypothesized to relate to failure to distinguish the internal voice and external sounds, and this deficit is thought to be associated with impaired self-prediction in comparing external and self-generated contents. In this magnetoencephalographic imaging (MEGI) study, we compared SIS M100 in the primary auditory cortex (A1) between the healthy controls (HC; N = 30) and SZ patients (N = 22). The SZ patients displayed reduced SIS and M100 in the A1, and this impairment is negatively correlated with auditory hallucinations. These outcomes suggest that the SZ patients' hallucinatory symptoms are caused by misattribution between the external and self-generated stimuli. We proposed that the weakened self-agency and neural oscillations may lead to this misattribution.

Introduction

In ordinary cases, during speech monitoring, it is thought that efference copy self-prediction mechanisms suppress the auditory cortical response to self-generated sounds, compared to listening to external speech (Chang et al., 2013; Ford & Mathalon, 2012; Kort et al., 2014; Sitek et al., 2013). Self-generated (and therefore highly predictable) sounds give rise to suppressed responses, thus allowing speakers to pay better attention to sounds in the external environment (Chang et al., 2013; Ford & Mathalon, 2012) indicative of a primordial biological basis for self-agency that is essential for normal interactions with outside reality. This process of suppression in the A1 is called speaker-induced suppression (SIS), which indicates the difference in activation of the A1 when the speaker listens to their own auditory feedback during active vocalization or passive listening. Previous studies using functional imaging such as ECoG and MEG revealed that the greatest SIS difference was found in the A1, and this difference fades away when furthering from the A1(Chang et al., 2013; Kort et al., 2014). Another study examined this concept in patients with Alzheimer's disease (AD), and AD patients displayed attenuated SIS and M100 (the peak of difference between speaking and listening around 100 ms after the onset of voice feedback)(Kim et al., 2023). This indicates that a reduced SIS reveals neural abnormalities. However, the MEG SIS of the M100 at the A1 in patients with SZ remains unknown.

A dysfunctional predictive mechanism would lead to incorrect predictions, causing the misattribution of self-generated actions as externally generated. Hallucinations,

especially auditory verbal hallucinations (AVH), are one of the top-ranked positive symptoms in patients with schizophrenia (SZ). Patients with SZ can demonstrate just such difficulties when self-generated actions are experienced as being of alien origin—the misperception of self-generated speech as an auditory hallucination. Both psychophysical and neuroimaging studies have suggested that self-monitoring may be dysfunctional in patients with schizophrenia(Ford et al., 2001; Frith et al., 2000). A large body of literature has revealed that the primary auditory cortex (A1) is activated when SZ patients experience AVH(Dierks et al., 1999; Jardri et al., 2011; van de Ven et al., 2005). In the meanwhile, other MEG studies found that the SZ patients displayed abnormal prefrontal activities (Jia et al., 2023; Jia et al., 2024) and they are associated with hallucinations in SZ patients(Haas et al., 2021). Using the methodology of N1 of EEG SIS, a previous study found that the SZ patients displayed a reduced SIS, the deficit was observed more in those with more severe AVH(Ford et al., 2007). This suggests that the lack of suppression is an indicator of blurred boundaries between the external and self-generated speech, which the internal stimuli were experienced as external, is the impaired process of efference copy producing the corollary discharge from an expected sensory outcome. However, this suppression was only observed at the FCz (midline of the brain) location with EEG, and this was tested at the pre-talk stage. Therefore, a better functional neuroimaging study (i.e., MEG) focusing on the A1 SIS, and a correlation between this neural biomarker and AVH in SZ patients are needed to be investigated.

Here we directly test the hypothesis that patients with schizophrenia have a defect in their ability to predict the sensory consequences of their actions. We used MEG to investigate the cortical neural responses at speech onset (to examine SIS of M100) during the phonation of a single vowel in HC and patients with SZ. We correlated SIS with their AVH scores using two clinical assessments (SAPS and AVHRS). We hypothesized that the SZ patients would: (1) display reduced SIS and M100 compared to HC, and (2) negatively correlate this impairment with AVH scores. This would suggest that the SZ patients are impaired in efference copy self-prediction processes, and this failure might lead to other functional impairments in psychotic disorders.

Methods

Participants

The present study encompasses the baseline MEG portion of an NIMH-funded R01 (R01MH122897) study in schizophrenia. Without revealing the subjects' identities, 30 healthy controls (HC; comprising 20 males and 10 females, mean age = 43 years, mean education =17 years) and 23 schizophrenic patients (SZ; comprising 17 males, 5 females, and 1 non-binary, mean age = 36 years, mean education = 15 years) participated in the MEG study at the University of California San Francisco (UCSF). The study was approved by the Internal Review Board (IRB) at UCSF. All procedures were performed in accordance with NIMH and IRB guidelines at UCSF. Recruitment took place through our clinicaltrials.gov site (NCT04807530) or from our prior research studies if they had provided consent to be contacted for future studies. Participants were evaluated by a clinical psychologist and underwent clinical assessments to meet the inclusion criteria. Inclusion criteria for healthy participants were the absence of

neurologic psychiatric disorders (Axis I or Axis II (SCID-Nonpatient edition)) and major illnesses, no current or history of substance dependence or abuse, meets MRI criteria, good general physical health, age between 18 and 64 years, right-handed, and English as the first language. All participants provided written informed consent for this study and then completed structural magnetic resonance imaging (MRI) and the MEG speech-monitoring task at baseline. It is worth noting that one SZ participant was excluded for not being able to finish the clinical assessments attentively and patiently. Statistical analyses comparing the demographic characteristics between the HC and SZ groups were carried out using SPSS (IBM). See Table 1 for participants' demographics.

Clinical and Neuropsychological Assessments

Each SZ patient underwent a structured interview to determine the scores of the following assessments:

Scales for the Assessment of Positive Symptoms (SAPS). SAPS is an assessment composed of a list of positive symptom sections including hallucinations, delusions, bizarre behaviors, and thought disorders(Andreasen, 1984). The assessment has a total of 34 self-report questions, each rates the symptom severity from 0 (none) to 5 (severe). This study only focused on the hallucination section, which encompasses domains of auditory, olfactory, and visual hallucinations. A total of 7 questions in this section (including global hallucination rating) were evaluated and analyzed.

Auditory Vocal Hallucination Rating Scale (AVHRS). The AVHRS is a standardized interview with 16 questions designed to assess a patient's experiences over a set timeframe(Bartels-Velthuis et al., 2012). Besides the first question openly asking about

the number of voices in the past 30 days, responses were rated on a 5-point scale, ranging from 0 (none) to 4 (severe), aiding in determining the severity of symptoms. All questions were evaluated and analyzed.

Speech Monitoring Task

Participants completed a speech monitoring task in the MEG scanner at baseline and immediately after the rTMS session. They wore a microphone (AKG Pro Audio C520 Professional Head-Worn Condenser Microphone, AKG Acoustics, Vienna, Austria) and a pair of headphones. The microphone was connected to an amplifier, which, in turn, was connected to a Dell computer sound card (M-Audio Delta 44 4x4 analog I/O, M-Audio, Cumberland, RI). The amplified audio signal was transmitted back to participants through headphones. Before the experiment, participants confirmed that they could clearly hear the audio signal.

There were two sessions, Speak and Listen, included in the speech monitoring task. This experiment was programmed in MATLAB, each session consisted of 9 runs constituting 15 trials per run, totaling 135 trials. Each trial began with a visual cue green dot that appeared on the MEG-compatible screen. For the Speak session, participants were instructed to vocalize the vowel /□/ when they saw the green dot. They continued the phonation for 2.5 seconds until the dot disappeared while listening to real-time auditory feedback from the headphones. In the Listen session, participants passively listened to their own recorded phonation playback from the Speak session. The visual cue green dot was still displayed in the Listen session, and the participants

were directed to listen to their playbacks while looking at the green dot on the screen. The inter-trial interval was 2.5s during which time participants viewed a blank screen. Participants were given a 5-second break after every run (15 trials).

MRI Acquisition

Structural T1-weighted MRI images were obtained for each participant using a 3-Tesla Siemens MRI scanner at the UCSF Neuroscience Imaging Center for MEG source space reconstruction. For each subject, high-resolution MRI was acquired using an MPRAGE sequence (160 1-mm slices, field of view = 256 mm, repetition time = 2300 ms, echo time = 2.98 ms). For each subject, the outline of the brain on the structural scans was extracted, and the segmented brain was treated as a volume conductor model for the source reconstruction described below.

MEG Data Acquisition

All participants underwent speech monitoring tasks in the MEG scanner at UCSF. Magnetic fields were recorded using a whole-head 275 axial gradiometer MEG system (MEG International Services Ltd., Coquitlam, BC, Canada) in a shielded room. The sampling rate was 1200 Hz with a 0.001-300 Hz bandpass filter. Three fiducial coils (nasion, left, and right preauricular points) were placed to localize the head position relative to the sensors. Three fiducial coils (nasion and left and right preauricular) were placed to localize the position of the head relative to the sensor array for co-registration of the MEG data with each individual's structural anatomical MRI. Head localization was performed at the beginning and end of each task block to register head position and to

measure head movement during the task. Third gradient noise correction filters were applied to the data and corrected for a direct current offset based on the whole trial. Noisy sensors and trials with artifacts (i.e., due to head movement, eye blinks, and saccades or sensor noise) were defined as magnetic flux exceeding 2.5 pT. epochs were rejected from further analysis if they contained artifacts.

MEG Analyses

Condition-specific blocks were amalgamated to create separate MEG datasets for Speak and Listen sessions for each participant. To counteract disruptions from distant magnetic fields, 29 reference sensors were utilized, employing a synthetic third-order gradiometer calculation(Vrba & Robinson, 2001; Weinberg et al., 1984). Furthermore, a dual signal subspace projection (DSSP) algorithm was implemented for the participants who had unremovable and metal dental works to alleviate artifacts induced by them in biomagnetic measurements(Cai et al., 2019; Sekihara et al., 2016). The MEG data underwent a 2 Hz high-pass filter to eliminate gradual fluctuations and precise annotation was applied at the onset of vocalization feedback. Trials were segmented from -300ms to +300ms around the onset of phonation feedback and underwent correction for direct-current (DC) offset and filtering within the frequency range of 2 to 150 Hz. Trial rejection ensued if MEG sensor channels exceeded a 1.5 pT threshold or if speech was detected during Listen trials, with meticulous manual verification of flagged artifacts. Subsequently, trials were averaged to produce a singular time series per condition, further organized into distinct sensor arrays for the left and right

hemispheres to capture auditory responses from each hemisphere. Co-registration of the MEG data with each individual's structural anatomical MRI was performed based on the nasion, left, and right preauricular fiducial coil positions utilizing the Neurodynamic Utility Toolbox for MEG (NUTMEG; available at http://nutmeg.berkeley.edu).

Bayesian covariance beamforming was implied to the averaged trial for left and right sensor array location from each subject, focused on the left (MNI x, y, z = [-54.3, -26.5, 11.6]) and right (MNI x, y, z = [54.4, -26.7, 11.7]) primary auditory cortices (A1)(Kort et al., 2014). The consequent source timeseries was converted into absolute Bayesian Covariance Beamformer (BCB) activity from the Champagne algorithm (Bhutada et al., 2022; Cai et al., 2021) for each of the 481 time points (from -100th ms to +300th ms, a total range of 400ms), which produced a timeseries of positive-going evoked activity within the voxel most adjacent to each hemisphere's A1 MNI coordinate. Following this step, the extracted latencies for each session (Speak or Listen) and hemisphere were visually confirmed and evaluated. Normalized latency in both Speak and Listen sessions was computed using the mean and standard deviation for the z-scores of each participant. The normalized latencies were extracted for further statistical analyses. To calculate M100 evoked latency, we used the maximal Listen latency during the window of 50 ms before and after 100th ms to subtract the corresponding Speak latency at the same time point for each individual from each hemisphere, then computed the mean of M100 for each group and hemisphere.

Group Analyses

The Wilcoxon rank-sum test was performed to compare the latencies (in z-scores) between the HC and SZ groups at each of the 481 time points in each session (Listen or Speak) within the same hemisphere. The speaker-induced suppression (SIS) was calculated as the contrast between Listen and Speak sessions by simply subtracting Speak from Listen (SIS = Listen – Speak) in the identical hemisphere. The same rank-sum test was utilized to compare the SIS difference between HC and SZ groups. The significant threshold was set to $p < 0.05$ for both comparisons.

One-way ANOVA was carried out to compare the between-group intra-hemispheric differences and a two-tailed paired-sample t-test was performed to compare within-group cross-hemispheric differences of M100. M100 in each hemisphere was a within-group factor, and the group membership was a between-group factor. A p-value less than 0.05 indicates a statistically significant difference. We hypothesized the SZ group would display a lower A1 response with SIS and M100 compared to the HC group.

To study the relationship between SIS and SZ patients' hallucinations, a Pearson correlation was performed. In line with our hypotheses, we anticipated a negative association between SIS and hallucinations. Specifically, we expected a negative correlation between SAPS hallucination and SIS latency z-scores at each time point, along with AVHRS scores and SIS z-scores. The significant level was set to $p < 0.05$ with a 95% confidence interval. Therefore, a significant and the most negative correlation coefficient of ±50 ms around the 100th ms after the onset of speech feedback (0th ms) was extracted. The reference selection of 100ms was accorded to the previous

studies, where the A1 was most activated in response to the onset of speech feedback around 100ms(Kim et al., 2023; Kort et al., 2014).

Result

During the Listen session, the left A1 activation of the SZ group was significantly lower than HC when they were passively listening to their phonation recorded from the Speak session around the 100th ms after the playback started. This significant range is from 90 to 110 ms. SZ patients also showed reduced left A1 activation during the Speak session, in which participants listened to their online playback during their phonation. The significant time ranges from 112 to 149 ms.

Consistent with our hypothesis, with the rank-sum test, the average evoked SIS (Listen – Speak) amplitude was significantly lower in the SZ group around 100 ms after the phonation feedback ($p < .05$). However, this difference was only observed in the left hemisphere (Figure 3A) and absent in the right hemisphere (Figure 3B). The significant time ranges from 88 to 106 ms. We compared the peaks of M100 (Max Listen around 100ms – correspondent Speak) between groups, which indicate the peak of SIS of each participant around 100 ms of phonation. In the left hemisphere, one-way ANOVA showed a significant M100 peak difference ($F(1, 50) = 5.092$, $p = .028$) between HC (Mean = 4.76, SD = 3.03) and SZ (Mean = 3.08, SD = 2.04), this difference was absent in the right hemisphere ($F(1, 50) = 0.638$, $p = .43$)), which HC showed an average of 3.49 (SD = 2.30) and SZ showed 3.46 (SD = 1.79) M100 peak latency. The HC group

displayed lateral effect with significantly different M100 peaks ($t(29) = 2.79$, $p = .009$) between the left (Mean = 4.76, SD = 3.03) and right (Mean = 3.49, SD = 2.30) hemispheres (Figure 3C). The lateral effect was not observed in the SZ group ($p > .05$).

We also compared the time to reach the M100 peak between groups and hemispheres. In the left hemisphere, it took an average of 106.43 ms (SD = 13.2) for HC and 102.5 ms (SD = 9.51) for SZ to reach the M100 peak; in the right hemisphere, HC spent 100 ms (SD = 13.41) and SZ spent 101.9 ms (SD = 9.74) on reaching to the M100 peak. There is no between-group difference or laterality effects in time to reach the M100 peak (all $p > .05$), see Figure 3D.

Another aim of this study is to investigate the association between SIS and clinical hallucination scores in the SZ group. A significant cutoff of $p < .05$, and a most negative correlation between SIS evoke response around the 100th ms and neuropsychological assessment scores in SZ was extracted. For the correlation between SIS evoke response and the number of SAPS (Figure 4A), the correlation coefficient was significant from 75.5 to 89.6 ms, and the most negative one was at 86.3^{th} ms ($r = -.51$, $p < .01$). For the correlation between SIS and AVHRS voice hallucination number (Figure 4B), the correlation coefficient was significant from 76.3 to 87.9 ms, and the most negative one was at 86.3^{th} ms ($r = -.53$, $p = 0.012$).

Discussion

While speaking, an automatic and unconscious process occurs in which speakers tend to monitor and compare their sensory feedback with feedback predictions from a copy of motor signals, which is called efference copy (Chang et al., 2013; Ford & Mathalon, 2012; Houde et al., 2002; Kort et al., 2014). A hypothesis is that the goal of such suppression of predicted self-generated stimuli is to increase the sensitivity of the sensory system to external stimuli (Crapse & Sommer, 2008; Ozker et al., 2024; Poulet & Hedwig, 2006; Poulet et al., 2002; Schneider et al., 2018). This process is underlying self-agency, which is the sense of being an owner of one's thoughts and actions. Self-agency facilitates successfully interacting with the external environment by correctly distinguishing external and self-generated contexts. Aligned with our hypotheses, we observed more SIS M100 with MEG imaging in HC than in SZ groups in the left hemisphere. Following the rationale that self-agency modulates the ability to differentiate the stimuli generated by the self and the external environment, the SZ participants exhibited an impaired self-agency, which led to a deficit in distinguishing the self and the external inputs. This finding converges with previous studies that found reduced SIS in SZ patients using EEG, particularly with an approach of event-related potential N1/N100 (Ford & Mathalon, 2004; Griffiths et al., 2023; Perez et al., 2012; Sitek et al., 2013) at the Cz position of EEG. This narrowed difference in A1 activations between Speak and Listen conditions in the SZ group suggests the misperception of internal and external stimuli, and this can contribute to difficulty differentiating the sources of sensory perceptions.

In the present study, the findings of reduced SIS in SZ were driven by its two components, which are reduced left A1 activations in Speak (Figure 2C) and Listen (Figure 2A) conditions around 100 ms after the playback onset. In the part of Listen, our results converge with previous studies in which SZ patients showed deprivation of A1 activity while passively listening to the external sound, either from speakers' self-produced vowel (Ford & Mathalon, 2004) or pure tone (Doucet et al., 2019), compared to the HC participants. This subnormal phenomenon suggests that the SZ patients misattributed the external-generated stimuli. In this study, we also found that the SZ group displayed reduced left A1 activation in the Speak session compared to the HC group. This outcome is contradicted by the studies that examined EEG SIS and found increased A1 responses (Ford & Mathalon, 2004) , or this phenomenon was not obvious (Griffiths et al., 2023) in the SZ patients during the speaking task. However, EEG has a long tradition of inaccuracy in localizing specific brain regions, and this contradiction might be caused by the failure of spatial registration. Using methods with high temporospatial resolution, such as electrocorticography (ECoG) or MEG, previous studies were able to examine the neural responses across multiple nodes next to the A1 during the same speech tasks (Chang et al., 2013; Kort et al., 2014). They found that some temporal regions had relatively smaller activation during the speech, for instance, the superior/middle temporal gyrus (STG/MTG) in an MEG study (Kort et al., 2014) and the e22 in an ECoG study (Chang et al., 2013). This suggests that components of SIS are sensitive to spatial resolution, and the selection of interested region(s) and EEG studies may not be able to reveal a holistic dynamic of SIS. Our MEG study utilized a

specific MNI coordinate to locate the A1 in each hemisphere, and this is reasonable we revealed this reduction in the left A1 in SZ patients during the Speak session.

The current study is the first to show the impaired asymmetric laterality effect of A1 SIS in SZ patients. The left A1 of SZ patients was universally suppressed during Speak and Listen sessions, which directly led to impaired M100 SIS in the left hemisphere. This phenomenon is absent in the right hemisphere, although the right A1 displayed some levels of nonsignificant suppression (Figure 2B) compared to the HC group. A large body of prior studies has shown predominant left A1 activation in hearing tasks from vocalization (Binder et al., 2000; Curio et al., 2000; Houde et al., 2002; Kort et al., 2014; Zatorre et al., 2002; Zatorre et al., 1992). It is thought that the audition of vocalization has a solid relation to speech production, and interactions between these cortices are essential to successfully produce accurate speech (Parker et al., 2005). In most people, including the left-handed ones, who have their Broca's area in the left hemisphere (Knecht et al., 2000), the dominance of left A1 is reasonable. Therefore, the left A1 is supposed to exhibit predominant activation in response to vocalization.

However, the A1 in SZ is not always deactivated, especially in cases of absent external stimuli. Many functional studies have reported that patients with SZ exhibited A1 and Broca’s area activations during auditory verbal hallucinations (Dierks et al., 1999; Jardri et al., 2011; van de Ven et al., 2005). This might be caused by reduced A1 grey matter volume and increased connectivity between A1 and Broca’s area (Okuneye et al., 2020). This suggests that during auditory verbal hallucinations, the self-prediction system in SZ

patients misattributes the internal sound to external stimuli, which makes it difficult for the patients to distinguish external and self-generated stimuli. This inference converges with our out that SIS is negatively correlated with hallucination severity (Figure 4). Reduced SIS indicates that SZ patients do not distinguish the source of auditory stimuli and are more likely to misattribute and misperceive one to the other. This impairment is related to auditory hallucinations, in which SZ patients perceive the internal voice as an external sound and have difficulty successfully separating them.

This self-prediction mechanism is underlying self-agency, which is the awareness of being the agent of one's thoughts and actions. Many brain regions, such as the medial prefrontal cortex (mPFC), the paracentral lobule/supplementary motor area, the basal ganglia, and the cerebellum, are involved in experiencing self-agency (Constantinidis & Steinmetz, 2005; Cunnington et al., 2002; Khalighinejad et al., 2018; Welniarz et al., 2021). However, our previous study in healthy participants showed that self-agency can be enhanced by stimulating the mPFC (Tan et al., 2024). Moreover, the mPFC is involved in an inhibitory process (Knight et al., 1999), reduction of functional connectivity (Alonso-Solís et al., 2015; Haas et al., 2021; Jia et al., 2023) and grey matter volume (Song et al., 2015) in the mPFC are thought to be related to the failure suppression of auditory hallucination in A1. Therefore, future studies should focus on examining the causal relation between enhanced mPFC activity, self-agency, and its modulation of hallucinations in patients with SZ.

Limitations

There are three limitations to this study. First, although we set the maximal threshold of CPZ usage, patients' usage of antipsychotics varies, and we do not know how this will contribute to the overall brain activity and if the reduction of SIS is related to drug use. Second, SIS is only correlated with SZ patients' AVHRS number of hallucinations and overall SAPS hallucinations, but not other domains in AVHRS. This might be caused by the characteristics of patients since we only recruited patients with high functions to ensure the patients could complete the whole experiment. Lower positive symptoms might be due to this in the recruitment processes. The third limitation is that this study does not have enough HC and SZ participants, and therefore we are continuously recruiting more participants for the study

References

Alonso-Solís, A., Vives-Gilabert, Y., Grasa, E., Portella, M. J., Rabella, M., Sauras, R. B.,…Corripio, I. (2015). Resting-state functional connectivity alterations in the default network of schizophrenia patients with persistent auditory verbal hallucinations. *Schizophr Res*, *161*(2-3), 261-268. https://doi.org/10.1016/j.schres.2014.10.047

Andreasen, N. C. (1984). SCALE FOR THE ASSESSMENT OF POSITIVE SYMPTOMS. In.

Bartels-Velthuis, A. A., van de Willige, G., Jenner, J. A., & Wiersma, D. (2012). Consistency and reliability of the auditory vocal hallucination rating scale (AVHRS). *Epidemiol Psychiatr Sci*, *21*(3), 305-310. https://doi.org/10.1017/S2045796012000108

Bhutada, A. S., Cai, C., Mizuiri, D., Findlay, A., Chen, J., Tay, A.,…Nagarajan, S. S. (2022). Clinical Validation of the Champagne Algorithm for Evoked Response Source Localization in Magnetoencephalography. *Brain Topography*, *35*(1), 96-107. https://doi.org/10.1007/s10548-021-00850-4

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Springer, J. A., Kaufman, J. N., & Possing, E. T. (2000). Human Temporal Lobe Activation by Speech and Nonspeech Sounds. *Cerebral Cortex*, *10*(5). https://doi.org/10.1093/cercor/10.5.512

Cai, C., Chen, J., Findlay, A. M., Mizuiri, D., Sekihara, K., Kirsch, H. E., & Nagarajan, S. S. (2021). Clinical Validation of the Champagne Algorithm for Epilepsy Spike Localization. *Frontiers in Human Neuroscience*, *15*.

Cai, C., Kang, H., Kirsch, H. E., Mizuiri, D., Chen, J., Bhutada, A.,…Nagarajan, S. S. (2019). Comparison of DSSP and tSSS algorithms for removing artifacts from vagus nerve stimulators in magnetoencephalography data. *Journal of Neural Engineering*, *16*(6), 066045. https://doi.org/10.1088/1741-2552/ab4065

Chang, E. F., Niziolek, C. A., Knight, R. T., Nagarajan, S. S., & Houde, J. F. (2013). Human cortical sensorimotor network underlying feedback control of vocal pitch. *Proc Natl Acad Sci U S A*, *110*(7), 2653-2658. https://doi.org/10.1073/pnas.1216827110

Constantinidis, C., & Steinmetz, M. A. (2005). Posterior Parietal Cortex Automatically Encodes the Location of Salient Stimuli. *Journal of Neuroscience*, *25*(1). https://doi.org/10.1523/JNEUROSCI.3379-04.2005

Crapse, T. B., & Sommer, M. A. (2008). Corollary discharge across the animal kingdom. *Nat Rev Neurosci*, *9*(8), 587-600. https://doi.org/10.1038/nrn2457

Cunnington, R., Windischberger, C., Deecke, L., & Moser, E. (2002). The preparation and execution of self-initiated and externally-triggered movement: a study of event-related fMRI. *Neuroimage*, *15*(2), 373-385. https://doi.org/10.1006/nimg.2001.0976

Curio, G., Neuloh, G., Numminen, J., Jousmäki, V., & Hari, R. (2000). Speaking modifies voice-evoked activity in the human auditory cortex. *Human Brain Mapping*, *9*(4), 183-191. https://doi.org/10.1002/(SICI)1097-0193(200004)9:4<183::AID-HBM1>3.0.CO;2-Z

Dierks, T., Linden, D. E. J., Jandl, M., Formisano, E., Goebel, R., Lanfermann, H., & Singer, W. (1999). Activation of Heschl's Gyrus during Auditory Hallucinations. *Neuron*, *22*(3). https://doi.org/10.1016/S0896-6273(00)80715-1

Doucet, G. E., Luber, M. J., Balchandani, P., Sommer, I. E., Frangou, S., Doucet, G. E.,…Frangou, S. (2019). Abnormal auditory tonotopy in patients with schizophrenia. *npj Schizophrenia 2019 5:1*, *5*(1). https://doi.org/10.1038/s41537-019-0084-x

Ford, J. M., & Mathalon, D. H. (2004). Electrophysiological evidence of corollary discharge dysfunction in schizophrenia during talking and thinking. *Journal of Psychiatric Research*, *38*(1), 37-46. https://doi.org/10.1016/S0022-3956(03)00095-5

Ford, J. M., & Mathalon, D. H. (2012). Anticipating the future: automatic prediction failures in schizophrenia. *Int J Psychophysiol*, *83*(2), 232-239. https://doi.org/10.1016/j.ijpsycho.2011.09.004

Ford, J. M., Mathalon, D. H., Heinks, T., Kalba, S., Faustman, W. O., & Roth, W. T. (2001). Neurophysiological Evidence of Corollary Discharge Dysfunction in Schizophrenia. *American Journal of Psychiatry*, *158*(12). https://doi.org/10.1176/appi.ajp.158.12.2069

Ford, J. M., Roach, B. J., Faustman, W. O., & Mathalon, D. H. (2007). Synch Before You Speak: Auditory Hallucinations in Schizophrenia. *American Journal of Psychiatry*, *164*(3), 458-466. https://doi.org/10.1176/ajp.2007.164.3.458

Frith, C. D., Blakemore, S., & Wolpert, D. M. (2000). Explaining the symptoms of schizophrenia: abnormalities in the awareness of action. *Brain Res Brain Res Rev*, *31*(2-3), 357-363. https://doi.org/10.1016/s0165-0173(99)00052-1

Griffiths, O., Jack, B. N., Pearson, D., Elijah, R., Mifsud, N., Han, N.,…Whitford, T. J. (2023). Disrupted auditory N1, theta power and coherence suppression to willed speech in people with schizophrenia. *NeuroImage : Clinical*, *37*. https://doi.org/10.1016/j.nicl.2022.103290

Haas, S. S., Hinkley, L. B. N., Fisher, M., Vinogradov, S., Nagarajan, S., & Subramaniam, K. (2021). A Neural Biomarker for Hallucinations: Medial Prefrontal Aberrations in Neural Connectivity Predict Self-Agency Deficits and Hallucination Severity in Schizophrenia. *J Brain Res*, *4*(3).

Houde, J. F., Nagarajan, S. S., Sekihara, K., & Merzenich, M. M. (2002). Modulation of the Auditory Cortex during Speech: An MEG Study. *Journal of Cognitive Neuroscience*, *14*(8), 1125-1138. https://doi.org/10.1162/089892902760807140

Jardri, R., Pouchet, A., Pins, D., & Thomas, P. (2011). Cortical Activations During Auditory Verbal Hallucinations in Schizophrenia: A Coordinate-Based Meta-Analysis. *American Journal of Psychiatry*, *168*(1). https://doi.org/10.1176/appi.ajp.2010.09101522

Jia, Y., Jariwala, N., Hinkley, L. B. N., Nagarajan, S., & Subramaniam, K. (2023). Abnormal resting-state functional connectivity underlies cognitive and clinical symptoms in patients with schizophrenia. *Frontiers in Human Neuroscience*, *17*. https://doi.org/10.3389/fnhum.2023.1077923

Jia, Y., Kudo, K., Jariwala, N., Tarapore, P., Nagarajan, S., & Subramaniam, K. (2024). Causal role of medial superior frontal cortex on enhancing neural information flow and self-agency judgments in the self-agency network. *medRxiv*. https://doi.org/10.1101/2024.02.13.24302764

Khalighinejad, N., Schurger, A., Desantis, A., Zmigrod, L., & Haggard, P. (2018). Precursor processes of human self-initiated action. *Neuroimage*, *165*, 35-47. https://doi.org/10.1016/j.neuroimage.2017.09.057

Kim, K. X., Dale, C. L., Ranasinghe, K. G., Kothare, H., Beagle, A. J., Lerner, H.,…Houde, J. F. (2023). Impaired Speaking-Induced Suppression in Alzheimer's Disease. *eNeuro*, *10*(6). https://doi.org/10.1523/ENEURO.0056-23.2023

Knecht, S., Dräger, B., Deppe, M., Bobe, L., Lohmann, H., Flöel, A.,…Henningsen, H. (2000). Handedness and hemispheric language dominance in healthy humans. *Brain*, *123*(12). https://doi.org/10.1093/brain/123.12.2512

Knight, R. T., Staines, W. R., Swick, D., & Chao, L. L. (1999). Prefrontal cortex regulates inhibition and excitation in distributed neural networks. *Acta Psychol (Amst)*, *101*(2-3), 159-178. https://doi.org/10.1016/s0001-6918(99)00004-9

Kort, N. S., Nagarajan, S. S., & Houde, J. F. (2014). A bilateral cortical network responds to pitch perturbations in speech feedback. *NeuroImage*, *86*, 525-535. https://doi.org/10.1016/j.neuroimage.2013.09.042

Okuneye, V. T., Meda, S., Pearlson, G. D., Clementz, B. A., Keshavan, M. S., Tamminga, C. A.,…Keedy, S. K. (2020). Resting state auditory-language cortex connectivity is associated with hallucinations in clinical and biological subtypes of psychotic disorders. *Neuroimage Clin*, *27*, 102358. https://doi.org/10.1016/j.nicl.2020.102358

Ozker, M., Yu, L., Dugan, P., Doyle, W., Friedman, D., Devinsky, O., & Flinker, A. (2024). Speech-induced suppression and vocal feedback sensitivity in human cortex. *eLife*, *13*. https://doi.org/10.7554/eLife.94198.1

Parker, G. J., Luzzi, S., Alexander, D. C., Wheeler-Kingshott, C. A., Ciccarelli, O., & Lambon Ralph, M. A. (2005). Lateralization of ventral and dorsal auditory-language pathways in the human brain. *Neuroimage*, *24*(3), 656-666. https://doi.org/10.1016/j.neuroimage.2004.08.047

Perez, V. B., Ford, J. M., Roach, B. J., Loewy, R. L., Stuart, B. K., Vinogradov, S., & Mathalon, D. H. (2012). Auditory Cortex Responsiveness During Talking and Listening: Early Illness Schizophrenia and Patients at Clinical High-Risk for Psychosis. *Schizophrenia Bulletin*, *38*(6), 1216-1224. https://doi.org/10.1093/schbul/sbr124

Poulet, J. F. A., & Hedwig, B. (2006). The Cellular Basis of a Corollary Discharge. *Science*, *311*(5760). https://doi.org/10.1126/science.1120847

Poulet, J. F. A., Hedwig, B., Poulet, J. F. A., & Hedwig, B. (2002). A corollary discharge maintains auditory sensitivity during sound production. *Nature 2002 418:6900*, *418*(6900). https://doi.org/10.1038/nature00919

Schneider, D. M., Mooney, R., Schneider, D. M., & Mooney, R. (2018). How Movement Modulates Hearing. *Annual Review of Neuroscience*, *41*(Volume 41, 2018). https://doi.org/10.1146/annurev-neuro-072116-031215

Sekihara, K., Kawabata, Y., Ushio, S., Sumiya, S., Kawabata, S., Adachi, Y., & Nagarajan, S. S. (2016). Dual signal subspace projection (DSSP): a novel algorithm for removing large interference in biomagnetic measurements. *Journal of Neural Engineering*, *13*(3), 036007. https://doi.org/10.1088/1741-2560/13/3/036007

Sitek, K. R., Mathalon, D. H., Roach, B. J., Houde, J. F., Niziolek, C. A., & Ford, J. M. (2013). Auditory cortex processes variation in our own speech. *PLoS One*, *8*(12), e82925. https://doi.org/10.1371/journal.pone.0082925

Song, J., Han, D. H., Kim, S. M., Hong, J. S., Min, K. J., Cheong, J. H., & Kim, B. N. (2015). Differences in gray matter volume corresponding to delusion and hallucination in patients with schizophrenia compared with patients who have bipolar disorder. *Neuropsychiatric Disease and Treatment*, *11*. https://doi.org/10.2147/NDT.S80438

Tan, S., Jia, Y., Jariwala, N., Zhang, Z., Brent, K., Houde, J.,…Subramaniam, K. (2024). A randomised controlled trial investigating the causal role of the medial prefrontal cortex in

mediating self-agency during speech monitoring and reality monitoring. *Sci Rep*, *14*(1), 5108. https://doi.org/10.1038/s41598-024-55275-3

van de Ven, V. G., Formisano, E., Röder, C. H., Prvulovic, D., Bittner, R. A., Dietz, M. G.,…Linden, D. E. (2005). The spatiotemporal pattern of auditory cortical responses during verbal hallucinations. *Neuroimage*, *27*(3), 644-655. https://doi.org/10.1016/j.neuroimage.2005.04.041

Vrba, J., & Robinson, S. E. (2001). Signal Processing in Magnetoencephalography. *Methods*, *25*(2), 249-271. https://doi.org/10.1006/meth.2001.1238

Weinberg, H., Brickett, P. A., Vrba, J., Fife, A. A., & Burbank, M. B. (1984). The Use of a Squid Third Order Spatial Gradiometer to Measure Magnetic Fields of the Braina. *Annals of the New York Academy of Sciences*, *425*(1), 743-752. https://doi.org/10.1111/j.1749-6632.1984.tb23597.x

Welniarz, Q., Worbe, Y., & Gallea, C. (2021). The Forward Model: A Unifying Theory for the Role of the Cerebellum in Motor Control and Sense of Agency. *Frontiers in Systems Neuroscience*, *15*. https://doi.org/10.3389/fnsys.2021.644059

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends Cogn Sci*, *6*(1), 37-46. https://doi.org/10.1016/s1364-6613(00)01816-7

Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of Phonetic and Pitch Discrimination in Speech Processing. *Science*, *256*(5058). https://doi.org/10.1126/science.256.5058.846

Table 1. Subject characteristics (Mean, SD)

	HC (N = 30)	SZ (N = 22)	P-value
Age (years)	43 (17)	36 (11)	.107
Gender	20M, 10F	17M, 4F, 1NB	.35
Education (years)	17 (2.2)	15 (2.1)	.93
PANSS Positive	N/A	30 (9.2)	N/A
PANSS Negative	N/A	14 (4.0)	N/A
AVHRS Total	N/A	20.5 (15.0)	N/A
CPZ Equivalent	N/A	165 (250)	N/A
Illness Duration (years)	N/A	10 (9.6)	N/A

Note. NB = non-binary. PANSS = positive and negative syndrome scale. AVHRS = auditory vocal hallucination rating scale. CPZ = Chlorpromazine.

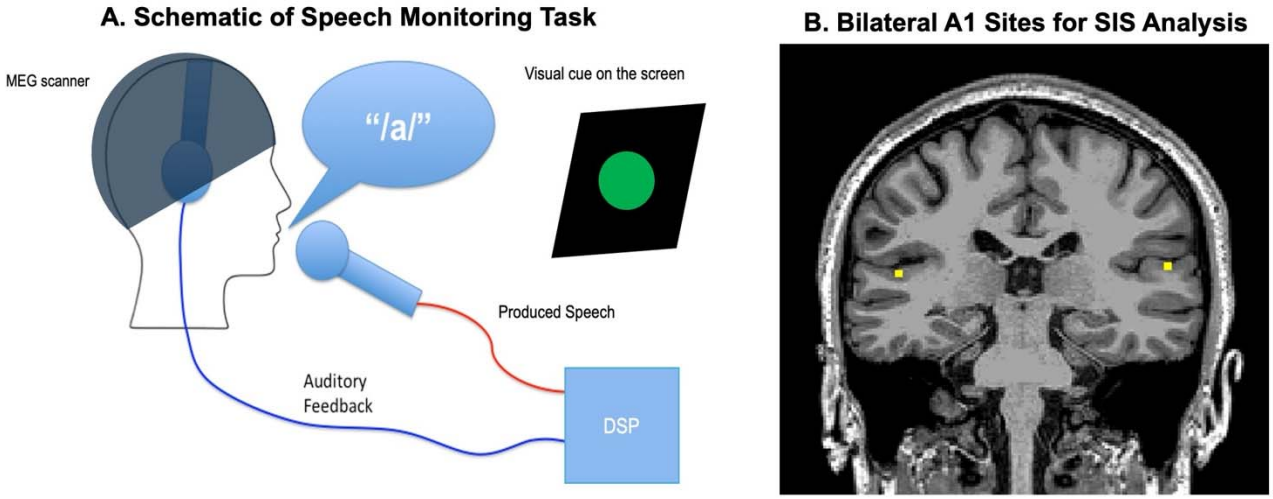

Figure 1. A. During the MEG speaking portion of the task, participants were instructed to vocalize the vowel /□/ when they saw the green dot while listening to real-time auditory feedback from the headphones via the digital signal processing (DSP) program. During the MEG listen session, participants passively listened to their own recorded phonation playback from the speak session. **B.** Anatomical coordinates based on each participant's structural MRI in each hemisphere, as shown by the yellow voxel ($8mm^3$) used to select the M100 auditory evoked potential from the primary auditory cortex (A1) for MEG SIS analyses (LH MNI: x, y, z -54, -27, 12; RH MNI: x, y, z = 54, -27,12).

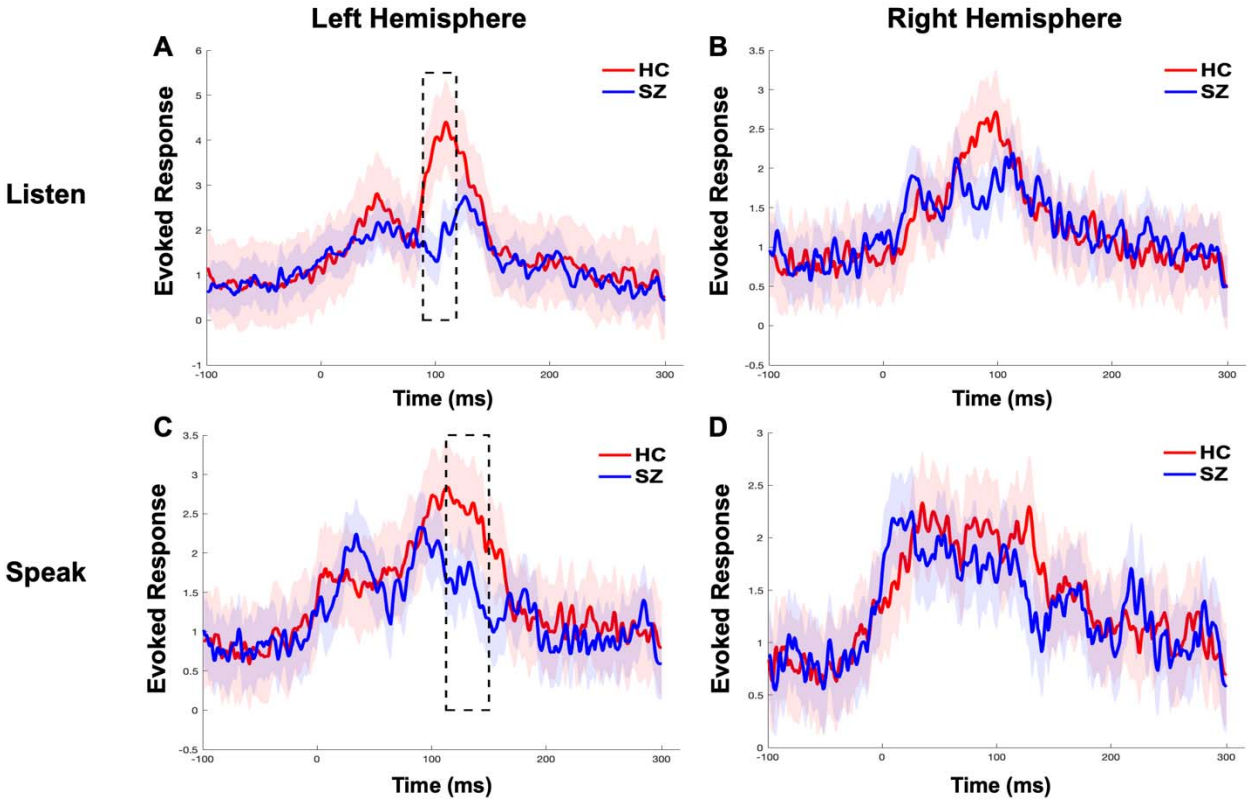

Figure 2. M100 responses indicating neural responses from the primary auditory cortex (A1), reveal that HC participants show significantly enhanced only left hemisphere auditory responses during the listen condition (A) and speak condition (C), compared to SZ.

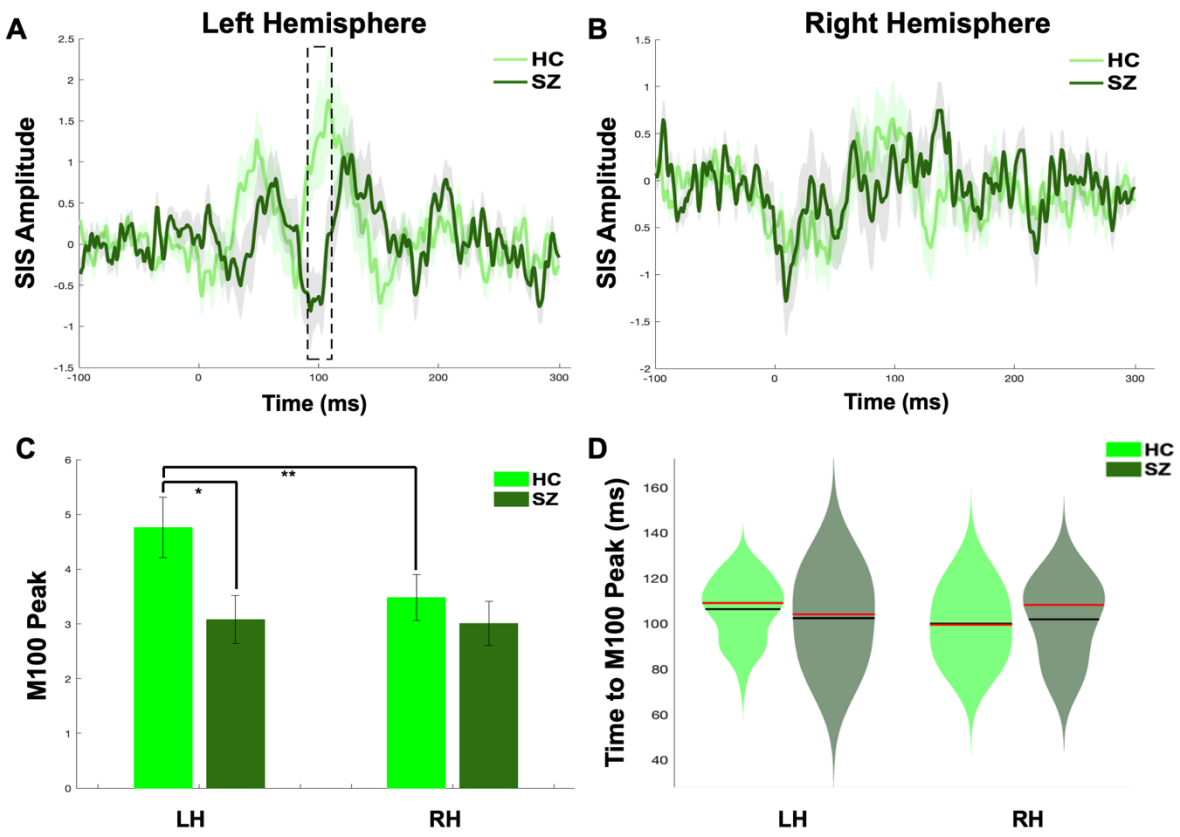

Figure 3. A-B.SIS peak amplitude (z-scores) in the left and right hemispheres was calculated by subtracting the M100 auditory evoked response during the speak condition from the listen peak amplitude (i.e., SIS Amplitude = Listen – Speak). The black dashed box indicates the time window (90-105ms) that showed a significant SIS difference only in the left hemisphere between HC and SZ groups ($p < .05$). 0 ms indicates the voice feedback onset in speak and listen sessions. **C**. Bar charts illustrate a laterality effect in which HC reveals a significantly larger M100 peak SIS amplitude (i.e., Listen-Speak z-scores) in the left hemisphere compared to the right hemisphere (**$p<.01$). This laterality effect is absent in SZ. HC also shows a significantly greater M100 SIS response in the left hemisphere, compared to SZ (*$p< .05$). **D**. The violin chart shows that there is no between-group difference or laterality effects in time to reach the M100 peak (median=red line; mean=black line).

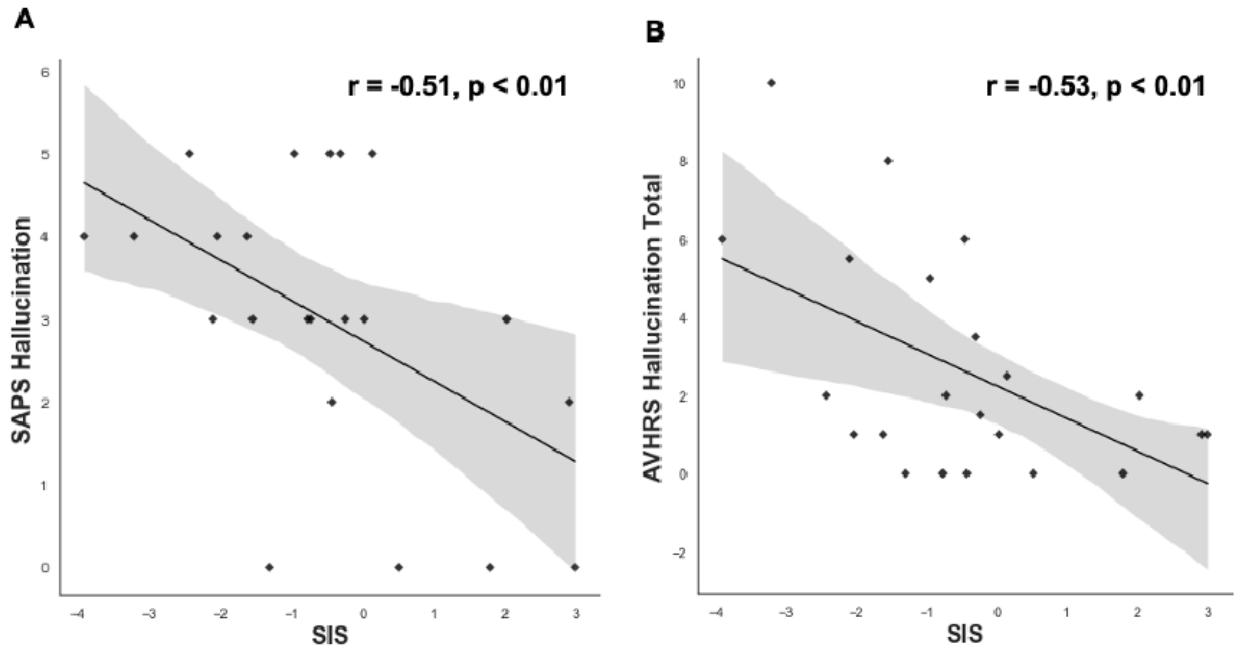

Figure 4. The scatter plots illustrate a significant negative correlation between SIS magnitude and hallucination severity, assessed with two independent scales, shown by the **(A)** SAPS hallucination scale, and **(B)** the AVHRS hallucination score. The shaded area indicates 95% confidence intervals.