

SCIENTIFIC REPORTS



OPEN

Voice-selective prediction alterations in nonclinical voice hearers

Ana P. Pinheiro^{1,2} , Michael Schwartz³ & Sonja A. Kotz^{3,4}

Auditory verbal hallucinations (AVH) are a cardinal symptom of psychosis but also occur in 6–13% of the general population. Voice perception is thought to engage an internal forward model that generates predictions, preparing the auditory cortex for upcoming sensory feedback. Impaired processing of sensory feedback in vocalization seems to underlie the experience of AVH in psychosis, but whether this is the case in nonclinical voice hearers remains unclear. The current study used electroencephalography (EEG) to investigate whether and how hallucination predisposition (HP) modulates the internal forward model in response to self-initiated tones and self-voices. Participants varying in HP (based on the Launay-Slade Hallucination Scale) listened to self-generated and externally generated tones or self-voices. HP did not affect responses to self vs. externally generated tones. However, HP altered the processing of the self-generated voice: increased HP was associated with increased pre-stimulus alpha power and increased N1 response to the self-generated voice. HP did not affect the P2 response to voices. These findings confirm that both prediction and comparison of predicted and perceived feedback to a self-generated voice are altered in individuals with AVH predisposition. Specific alterations in the processing of self-generated vocalizations may establish a core feature of the psychosis continuum.

Why do some people hear voices in the absence of external stimulation? Self-voice processing relies in part on the capacity to predict the sensory consequences of self-generated vocalizations and on rapidly detecting a discrepancy between predicted and perceived sensations^{1,2}. Auditory cortex activity is suppressed in response to self-generated (i.e., fully predictable) compared to external (i.e., less predictable) voices^{3–8}. The suppression of sensory cortical responses to self-generated stimuli most likely reflects the tagging of sensations as self-produced to avoid confusion with sensations in the external environment^{5,9}. There is compelling evidence that auditory verbal hallucinations (AVH; the perception of voices in the absence of a voice) result from the inability to predict the sensory consequences of self-generated signals due to a dysfunctional internal forward model^{10,11}.

Neural Signatures of the Internal Model

Studies exploring the internal forward model in voice/sound processing have assessed auditory cortex activity by using one of two comparison strategies: listening to self-produced voice in speech production vs. listening to pre-recorded self-voice^{7,9} or listening to a sound triggered by a button press vs. passively listening to the same sound^{12–15}. As the latter paradigm overcomes artifacts related to self-generated voice production (bone conduction; motor activity), it provides some advantages over the monitoring of activity during speech production. Both strategies focus on neural correlates of two aspects of the internal forward model: the *efference copy* (a copy of the motor command that is sent to the auditory cortex via the cerebellum and supports the computation of predicted sensory consequences) and the *corollary discharge* (the expected sensory feedback resulting from a self-generated action)^{9,15–17}. Suppression of sensory cortical responses to self-generated stimuli is assumed to reflect the tagging of sensations as self-produced to avoid confusion with sensations originating from the external environment¹⁷.

Event-related potential (ERP) studies have highlighted the role of the auditory N1 as a putative measure of sensory prediction. The N1 amplitude is suppressed in response to self-generated relative to externally generated voices^{4–8}. Similarly, a tone elicited by a button-press leads to an N1 suppression when compared to an externally

¹Faculdade de Psicologia, Universidade de Lisboa, Lisboa, Portugal. ²Neuropsychophysiology Lab, School of Psychology, University of Minho, Braga, Portugal. ³Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands. ⁴Department of Neuropsychology, Max-Planck-Institute for Human Cognitive and Brain Sciences, Leipzig, Germany. Correspondence and requests for materials should be addressed to A.P.P. (email: appinheiro@psicologia.ulisboa.pt)

presented tone^{12–15,18,19}. As the auditory N1 is primarily generated in the primary and secondary auditory cortices^{20–22}, these findings suggest that the suppression effect associated with the efference copy reflects reduced activity in these brain regions. When auditory feedback does not match the predicted sensation, a prediction error signal is generated, resulting in an increased N1 response^{3,15}. Larger N1 responses may signal increased attention to an unexpected sensory event (prediction error)^{23–25}.

Sensory-related suppression effects have also been observed in the P2 component of the ERP^{3,26}. Whereas N1 suppression is typically directly associated with the expected sensory feedback, the P2 may indicate a more conscious distinction between self-generated and externally-generated sensory events^{15,18,19}. However, P2 suppression is not consistently reported. Whereas P2 suppression in response to self-generated tones was observed in button-press tasks (for example^{15,18,19}), lack of P2 suppression to self-generated speech sounds in talking tasks has been proposed as a mechanism that allows preservation of the sensory experience of voice feedback during speech generation²⁷. These inconsistent findings indicate that the type of task may account for differences in P2 amplitude modulation.

Of note is that these ERP studies focus on neural responses obtained after the onset and during the processing of auditory feedback. However, patterns of neural activity preceding sensory feedback can shed light on the critical stage of sensory prediction formation *per se*. Specifically, pre-stimulus alpha activity is suggested to reflect the prediction of expected sensory consequences of an action^{28,29}. Pre-stimulus alpha power is enhanced in sensory cortices prior to self-induced speech²⁸, pure tones²⁹, or visual stimuli³⁰. Further, increased pre-stimulus alpha power for self-generated sounds is associated with larger N1 suppression²⁹. Hence, suppression may index the transfer of an efference copy of a motor command to the auditory cortex, whereas the N1 modulation may reflect how well the sensory consequences of an action match or mismatch (reflecting the magnitude of the prediction error), while pre-stimulus oscillatory power may reflect prediction *per se*²⁹.

Internal forward Model and Auditory Verbal Hallucinations

AVH are one of the cardinal symptoms of schizophrenia and experienced by up to 70% of schizophrenia patients³¹. However, they are also present in 6–13% of general population^{32,33}. The experience of AVH in psychotic and nonpsychotic individuals seems to engage similar cognitive mechanisms and brain areas^{34,35}. This suggests a neural substrate specific to AVH rather than schizophrenia. Nonetheless, despite numerous attempts to explain the neurofunctional mechanisms of AVH, they remain one of the notoriously unexplained symptoms in psychosis.

A substantial body of evidence shows that a failure to distinguish between internally and externally generated sensory signals (e.g., one's own voice *vs.* somebody else's voice) may underlie the experience of AVH^{34,36,37}. This is reflected in a reduced N1 suppression effect when listening to real-time feedback of one's own voice^{6,11,38} or in button-press tasks (contrasting self-initiation of a sound with passive exposure to the same sound)¹³. However, the later P2 ERP component as well as pre-stimulus oscillatory EEG activity have not been systematically analyzed in these studies. Further, as most of the existing studies involved chronic schizophrenia patients, it is possible that confounding effects associated with medication, hospitalization, and the presence of negative symptoms may have mediated the reported N1 results.

The study of nonpsychotic individuals who hear voices thus provides a necessary next step to understand the role of altered voice processing in AVH. These individuals are characterized by an increased tendency to falsely report the presence of a voice in bursts of noise³⁹ or in conditions of stimulus ambiguity in signal detection tasks³⁵. Moreover, they recognize words in degraded speech earlier than controls and before being explicitly informed of its intelligibility⁴⁰. Nonetheless, it remains unclear how these individuals process self-generated *vs.* externally-generated voice and sound feedback. If altered sensory feedback is observed as a function of increased hallucinatory predisposition, this would support the psychosis continuum hypothesis³¹, revealing common physiological brain processes underlying psychotic-like symptoms in nonpsychotic participants.

The Current Study and Hypotheses

The current study probed whether and how hallucination predisposition modulates the processing of fully predictable (i.e., self-generated) and less predictable (i.e., externally generated) auditory stimuli. By presenting both tones and voices after or in the absence of an action (button press), we examined whether potential problems with prediction formation (pre-stimulus alpha power) and sensory feedback (N1 and P2) are a characteristic in this sample of the general population, and whether the processing of simple (tones) and more complex (voice) auditory stimuli differ in this regard.

A well-established button-press paradigm^{15,18,19} was used. Next to investigating classical ERP components (N1 and P2), we performed a time-frequency analysis of pre-stimulus EEG activity in the alpha range (8–12 Hz). In line with the notion that the efficacy of an internal model is a crucial determinant of the experimental performance, these measures were taken as indices for three different processes: the formation of a prediction or efference copy (pre-stimulus alpha power); the comparison between predicted and perceived sensory feedback (N1); the conscious detection of a self-initiated sound (P2).

Our central hypothesis was that the experience of hallucinated voices involves alterations in the internal forward model. If psychotic-like experiences, such as AVH, are elicited by the same underlying neurocognitive mechanisms as in schizophrenia, nonpsychotic individuals with high hallucination predisposition (HP) should exhibit a similar pattern of EEG activity to that observed in schizophrenia. By forming an efference copy based on finger tapping, the auditory cortex will be prepared for incoming sensation and its activity will be attenuated in participants with lower HP (N1/P2 suppression) but not in those with higher HP¹³. Further, we hypothesized that pre-stimulus alpha power would be reduced for self-generated stimuli as a function of increased HP, reflecting altered prediction formation. We specifically hypothesized that altered prediction and sensory feedback in high HP would be enhanced for voice stimuli⁴¹.

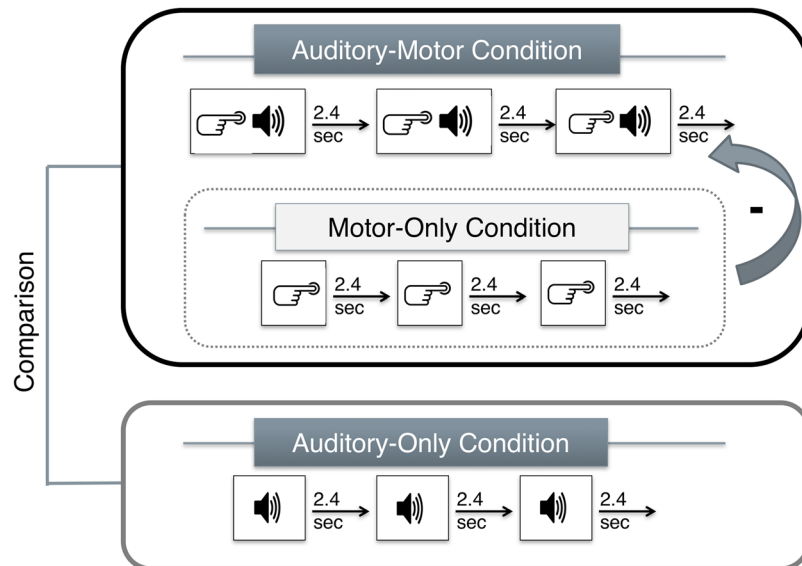


Figure 1. Schematic illustration of the three experimental blocks. A fixation cross was presented in the middle of a computer screen. Participants were instructed to change the index finger after 50 trials to ensure that the motor activation pattern was similar across participants during the AMC and AOC conditions, following indications on the screen (to avoid counting or cognitive demands associated with finger memorization). Change of index fingers was counterbalanced across participants. *Notes:* AMC = auditory-motor condition; AOC = auditory-only condition; MOC = motor-only condition. The AMC was corrected for motor activity based on a difference waveform (AMC-MOC). The statistical analysis involved the comparison of auditory ERPs elicited by the (corrected) AMC and the AOC.

Methods

Participants. The study involved two stages. In stage 1, a large sample of college students from different Universities in Portugal ($N = 354$) was recruited to complete an on-line version of the 16-items Launay-Slade Hallucination Scale (LSHS)⁴². The total score ranges between 0 and 64, with higher scores indicating higher hallucination predisposition. Responses are provided on a 5-point scale (0 = “certainly does not apply to me”; 1 = “possibly does not apply to me”; 2 = “unsure”; 3 = “possibly applies to me”; 4 = “certainly applies to me”). In stage 2, we recruited 49 participants from stage 1, who consented to be contacted for further research on voice processing. Over a 12-month recruitment period, 49 individuals were interviewed in more detail about their experiences and clinical history. All participants completed a thorough clinical assessment that established that for those who reported hearing ‘voices’ (and thus scored higher on the scale), voices were distinct from thoughts, were experienced at least once a month, were unrelated to drug or alcohol abuse, and that participants did not have a psychiatric diagnosis or received a psychiatric diagnosis in relation to voice-hearing.

From the 49 participants recruited, 32 participants varying in their LSHS scores (total score: $M = 22.72$, $SD = 14.03$, range = 0–51; auditory score: $M = 3.19$, $SD = 2.95$, range = 0–9) accepted to partake in an EEG experiment ($M_{age} = 22.77$, $SD = 4.06$, age range = 18–32 years; 18 females). Note that recruitment of nonpsychotic voice hearers represents a challenging process due to concerns about stigma as noted in previous studies⁴⁰. *A priori* power calculations using G*Power-3 statistical software⁴³ indicated that with a $\alpha = 0.05$ and power (1-error probability) = 0.90, a sample of a minimum 24 (based on a medium effect size of 0.25) participants would be required to allow the detection of such effects.

Participants were all right-handed⁴⁴, reported normal or corrected-to-normal visual acuity, and normal hearing. All participants provided informed consent and were reimbursed for their time, either by course credits or a voucher. The study was conducted in accordance with the Declaration of Helsinki and was approved by the local Ethics Committee of the University of Minho, Braga (Portugal).

Stimuli. A 680 Hz tone (50 ms duration; 70 dB sound pressure level [SPL]) and a pre-recorded self-voice speech sound (vowel/a/) were presented in Experiment 1 and 2, respectively. Before Experiment 2, a voice recording session took place: participants were instructed to repeatedly vocalize the syllable “ah”. Recordings were made with an Edirol R-09 recorder and CS-15 cardioid-type stereo microphone^{37,45}. After the recording, the best voice sample of the vowel /a/ from each participant (i.e., constant prosody; maximum duration of 300 ms) was selected. The voice sample was edited to eliminate background noise (using Audacity software), and a Praat script was applied to normalize intensity at 70 dB. The stimulus for each participant was saved as WAV file. Hence, all voice stimuli across participants had the same duration (300 ms) and intensity (70 dB SPL).

Procedure. In the EEG experiments, participants sat comfortably at a distance of 100 cm from a desktop computer monitor in a sound-attenuated and electrically shielded room. Each experiment included three conditions (see Fig. 1): auditory-motor (AMC), auditory-only (AOC), and motor-only (MOC)^{15,18,19}. In the AMC, a button

press instantaneously elicited a tone (Experiment 1) or the prerecorded voice of the participant (Experiment 2). Participants pressed a button approximately every 2.4 seconds. In the AOC, participants were instructed to passively listen and attend to the tones (Experiment 1) or to their pre-recorded self-voice (Experiment 2). The acoustic stimulation from the AMC was recorded on-line and used as the auditory sequence that was passively presented to participants in the AOC. The MOC represents a control condition that allowed controlling for motor-related artifacts via a button press (AMC-MOC): participants performed self-paced button presses approximately every 2.4 seconds but no tone (Experiment 1) or voice (Experiment 2) was elicited by the button press. No measurable sound was emitted by the presses. The AMC always preceded the AOC but the MOC was randomized across participants.

The experimental blocks were preceded by two training blocks^{15,18,19}. The training block contained 200 trials. Participants performed correct taps in 75% of trials. No feedback was provided during the experimental blocks. In each of the AMC and AOC, 100 trials were recorded. The MOC in each experiment consisted of 100 trials. Both experiments took place in the same EEG session, but their order was counterbalanced across participants.

The presentation and timing of the stimuli was controlled by Presentation software (version 16.3; Neurobehavioral Systems, Inc.). Auditory stimuli were presented via Sennheiser CX 300-II headphones. A BioSemi tapping device was used to record the finger taps.

EEG Data Acquisition and Analysis. EEG data were recorded using a 64-channel BioSemi Active Two system in a continuous mode at a digitization rate of 512 Hz and stored on hard disk for later analysis.

EEG data preprocessing details are presented as Supplementary Material. The ERP analysis followed Knolle *et al.*^{15,18,19}. The waveforms revealed two components: a negative component peaking at approximately 100 ms and a positive one peaking at approximately 200 ms (see Fig. 2). Peak amplitudes were calculated in the time windows of 70–110 ms for the N1 and 170–210 ms for the P2^{15,18,19}.

In EEG time-frequency analyses, power analysis was performed based on time-frequency magnitude values. The mean power in the alpha band (8–12 Hz) was calculated in a pre-stimulus interval (–250 to 0 ms)²⁹.

Amplitude and power data were extracted at four regions of Interest (ROI) comprising: left anterior (F7, F3, FT7, FC3, T7, C3), right anterior (F8, F4, FT8, FC4, T8, C4), left posterior (TP7, CP3, P7, P3, PO7, PO3), and right posterior (TP8, CP4, P8, P4, PO8, PO4) electrode positions^{15,18,19}.

Statistical Data Analyses. Amplitude and power data were analyzed with mixed linear models using the lmer⁴⁶ and lmerTest⁴⁷ packages in the R environment (R3.4.3. GUI 1.70), which were used to estimate fixed and random coefficients. In contrast to the more traditional repeated-measures ANOVA analysis, LMER allows controlling for the variance associated with random factors such as random effects for participants in ERP amplitude and EEG power measures⁴⁸. The default variance-covariance structure, i.e. the unstructured matrix, was used⁴⁹.

Results

Event-related-potentials. *N1.* Intra-class correlation coefficients indicated that 41% of the total variance in the N1 response was accounted for by differences between participants. A Gaussian distribution of residuals was selected to run the mixed model and probability plots (quantile-quantile plots) confirmed its adequacy. Amplitude was included as outcome, participants were included as random effects, whereas stimulus type (tone, voice), condition (self, external), ROI, and LSHS_{Total} were included as fixed effects.

The effects of stimulus type, condition, and ROI were significant. Specifically, we replicated the sensory N1 suppression effect: N1 was more negative for externally-generated compared to self-generated auditory stimuli ($\beta = -0.9498$, $SE = 0.3934$, $t = -2.414$, $p = 0.016$; see Fig. 2). Further, we observed that the N1 response was overall more negative for voices compared to tones ($\beta = -1.031$, $SE = 0.3934$, $t = -2.621$, $p = 0.009$). The typical frontocentral distribution of the N1^{50,51} was confirmed by the observation of less negative amplitude in the left posterior and right posterior ROIs compared to the left anterior ROI (*vs.* left posterior $-\beta = 1.4097$, $SE = 0.1676$, $t = 8.410$, $p < 0.001$; *vs.* right posterior $-\beta = 1.3882$, $SE = 0.1676$, $t = 8.282$, $p < 0.001$), and also to the right anterior ROI (*vs.* left posterior $-\beta = 1.6817$, $SE = 0.1653$, $t = 10.172$, $p < 0.001$; *vs.* right posterior $-\beta = 1.6602$, $SE = 0.1653$, $t = 10.042$, $p < 0.001$). The model identified a positive effect of LSHS_{Total} scores on the N1 amplitude, indicating that differences in hallucination predisposition accounted for N1 amplitude modulations ($\beta = 0.0361$, $SE = 0.0159$, $t = -2.621$, $p = 0.009$; see Fig. 2).

Based on our hypotheses, we tested whether LSHS_{Total} scores had a specific impact on the N1 amplitude as a function of stimulus type and condition, probing the interaction between the three predictors (see Table 1). For the sake of simplicity and facility in the interpretation, we defined the variable “interaction” (B) with 4 levels (2 stimulus types \times 2 conditions) and tested the following model in R: $m.N1 \leftarrow \text{lmer}(N1 \sim \text{LSHS_Total} * B + \text{ROI} + (1|\text{Subject}), \text{data} = N1, \text{REML} = \text{FALSE}); \text{summary}(m.N1)$, in which $N1\$B <- \text{interaction}(N1\$StimulusType, N1\$Condition)$. An increase in LSHS_{Total} scores was associated with a less negative N1 response for self-generated tones compared to self-generated voices ($\beta = 0.0423$, $SE = 0.0120$, $t = 3.535$, $p < 0.001$) as well as in a less negative N1 for externally-generated voices compared to self-generated voices ($\beta = 0.0301$, $SE = 0.0120$, $t = 2.517$, $p = 0.012$; see Fig. 3). Specifically, the N1 amplitude in response to the self-generated voice was expected to be more negative than the N1 amplitude for the externally-generated voice if $\text{LSHS_Total} > 18.94$, and the N1 amplitude for the self-generated voice was expected to be more negative than the amplitude for the self-generated tone if $\text{LSHS_Total} > 15.83$. A separate analysis of midline electrodes (considering anterior [Fz, Cz] *vs.* posterior [CPz, Pz] electrode sites) using the same statistical model described before revealed similar effects.

To probe the specific effect of AVH, we ran the same model replacing LSHS_{Total} with LSHS_{Auditory} (sum of items 4 – “In the past, I have had the experience of hearing a person’s voice and then found no one was there”, 8 – “I often hear a voice speaking my thoughts aloud”, and 9 – “I have been troubled by voices in my head”) as a fixed factor. The effect of LSHS_{Auditory} on the N1 amplitude was significant ($\beta = 0.1840$, $SE = 0.0849$, $t = 2.167$,

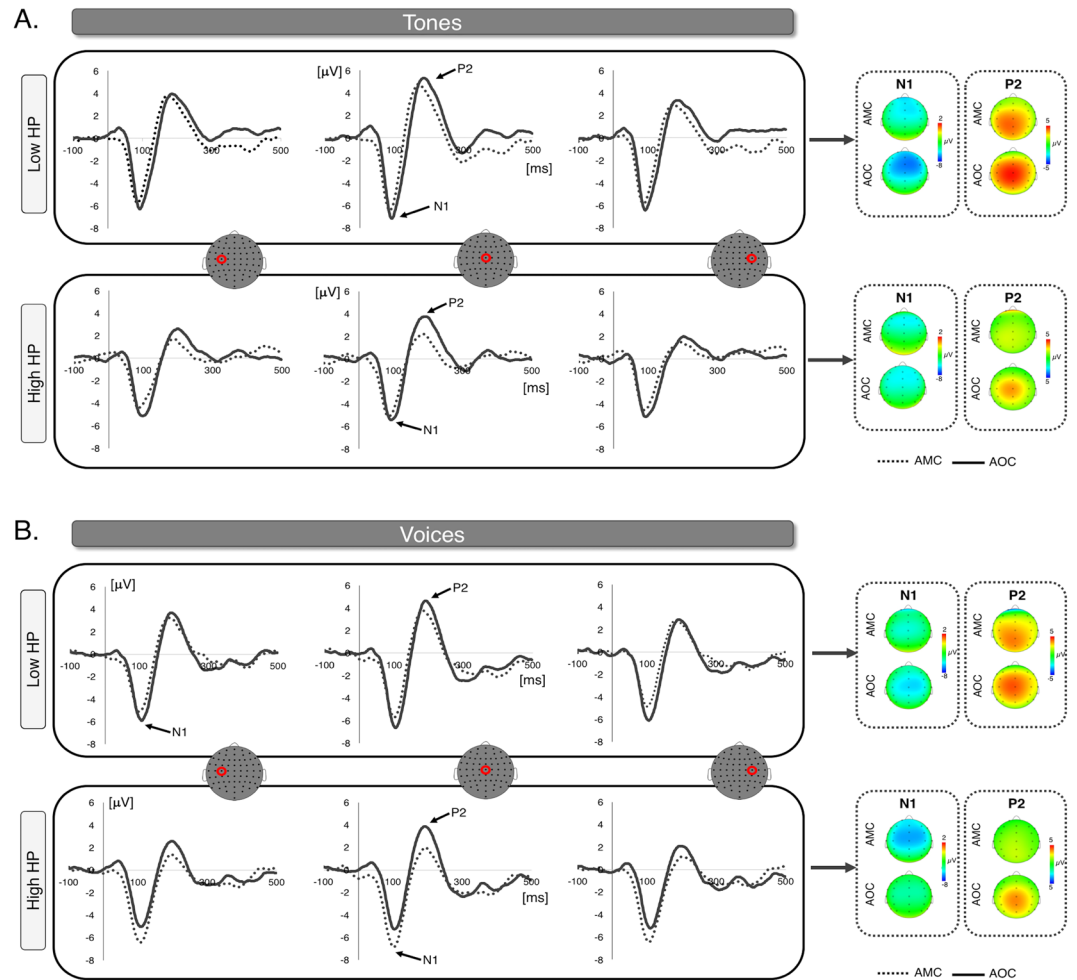


Figure 2. Grand average waveforms contrasting self-triggered and externally triggered tones (Panel A) and self-voices (Panel B) at electrodes C3, Cz and C4. A median split was performed on the data to illustrate differences in the processing of self- and externally triggered stimuli as a function of HP (high HP: ≥ 26 , $N = 17$, range $LSHS_{Total}$: 26–51; low HP: < 26 , $N = 15$, range $LSHS_{Total}$: 0–25). Topographic maps show voltage distribution in the 80–120 ms (N1) and 175–215 ms (P2) time windows. *Notes:* HP = hallucination predisposition; AMC = auditory-motor condition; AOC = auditory-only condition.

$p = 0.0352$). Further, the interaction of stimulus type and condition revealed a specific effect of the N1 amplitude modulation for the self-generated voice ($\beta = -0.1242$, $SE = 0.0600$, $t = -2.069$, $p = 0.039$).

P2. The analysis of P2 amplitude followed the same procedure as described above. Removing $LSHS_{Total}$ increased the goodness of fit of the model ($\chi^2(10) = 2.859$, $p < 0.001$), even though the Akaike's Information Criterion (AIC⁵²) for the complete model (AIC = 1849.9) was slightly lower than the AIC for the model without $LSHS_{Total}$ (AIC = 1850.8). Hence, $LSHS_{Total}$ was not included as a predictor of the P2 amplitude modulation.

The model showed that self-generated auditory stimuli were associated with a less positive P2 amplitude compared to externally-generated auditory stimuli ($\beta = 0.4626$, $SE = 0.1670$, $t = 2.770$, $p = 0.0058$; see Fig. 2). Stimulus type did not significantly predict the P2 amplitude modulation ($\beta = -0.2332$, $SE = 0.1670$, $t = -1.397$, $p = 0.1631$). A separate analysis of midline electrodes (considering anterior [Fz, Cz] vs. posterior [CPz, Pz] electrode sites) using the same statistical model described before revealed similar effects.

Pre-stimulus alpha power. The model including LSHS as a predictor had the best goodness of fit ($\chi^2(13) = 36.957$, $p < 0.001$; AIC with $LSHS_{Total} = -1691$, AIC without $LSHS_{Total} = -1662$; see Table 2). Alpha power was decreased before sounds that were externally-generated compared to self-generated sounds ($\beta = -0.0115$, $SE = 0.0054$, $t = -2.121$, $p = 0.035$; see Table 2 and Fig. 4).

HP specifically influenced alpha activity preceding the self-generated voice (see Table 2). An increase in $LSHS_{Total}$ resulted in a decrease in alpha power preceding self-generated tones compared to self-generated voices ($\beta = -0.0015$, $SE = 0.0003$, $t = -3.870$, $p < 0.001$) and preceding externally triggered voices compared to self-triggered voices ($\beta = 0.0017$, $SE = 0.0004$, $t = -4.756$, $p < 0.001$).

Variable	Estimate	SE	t-value	Pr(> z)
Fixed Effects				
Intercept	-5.44527	0.48904	-11.135	<0.001***
LSHS _{Total}	0.04619	0.01800	2.566	0.013612*
LSHS _{Total} * Self* Tone (vs. External Tone)	0.02261	0.011198	1.888	0.0596
LSHS _{Total} * External* Voice (vs. Self-voice)	0.030137	0.011976	2.517	0.012177
LSHS _{Total} * Self* Tone (vs. Self-Voice)	0.042339	0.011976	3.535	<0.001***
LSHS _{Total} * External* Voice (vs. External Tone)	0.01041	0.01198	0.869	0.3852
Groups				
	Name	Variance	SD	
Random Effects				
Subject	(Intercept)	1.538	1.240	
Residual		1.749	1.323	

Table 1. Linear mixed effects model of N1 amplitude including the effect of hallucination predisposition. *Notes.* SE = standard error; SD = standard deviation; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Degrees of freedom for Fixed Effects: $df = 480.00$ (except Intercept: $df = 50.01$).

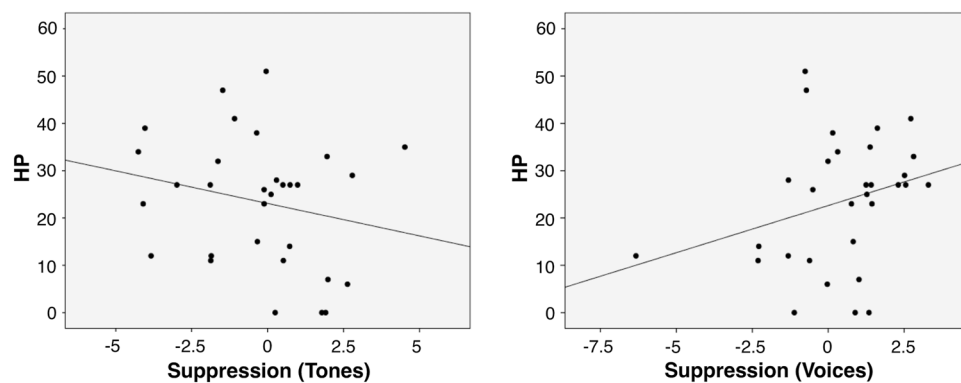


Figure 3. Relationship between hallucination predisposition (LSHS_{Total}) and N1 suppression to tones and voices. Note. Values represent difference in N1 amplitude between AOC and AMC conditions (negative values indicate less negative N1 in the AMC compared to AOC [AMC < AOC]; positive values indicate more negative N1 in the AMC compared to AOC [AMC > AOC]).

As for the N1, we probed the specific effect of AVH by replacing LSHS_{Total} with LSHS_{Auditory} as a fixed factor. The effect of LSHS_{Auditory} on pre-stimulus alpha power was not significant ($\beta = 0.0015$, $SE = 0.0023$, $t = 0.623$, $p = 0.538$).

Discussion

Why some people hear voices in the absence of external stimulation remains to be clarified. The current study probed how HP modulates the processing of auditory stimuli (tones vs. voices) that are anticipated as a consequence of a motor act (a button press) engaging an internal forward model. The study of nonpsychotic individuals with high HP represents an important step forward in probing the continuum hypothesis of psychosis. Here, we disentangled whether increased HP affects the formation of a prediction or efference copy (pre-stimulus alpha power), the comparison between predicted and perceived sensory feedback (N1), and/or the conscious detection of a self-initiated sound (P2). Our results suggest that an increase in HP is specifically related to alterations both in the generation of an efference copy and in the comparison between predicted and perceived sensory feedback. Further, they show that these alterations are more pronounced during the perception of one's own voice compared to simple tones.

The N1 suppression effect is reversed in high HP. Our findings replicate the classical sensory suppression effect (for example^{5-12,15,18,19,53}) both for tones and voices: self-generated sounds elicited a smaller N1 compared to externally generated sounds irrespective of stimulus complexity, an effect that was more prominent at anterior electrode sites^{50,51}. As the sensory consequences of self-initiated sounds are precisely predicted such that the auditory cortex is prepared to receive sensory feedback, the N1 amplitude is suppressed. We also observed that self-voices elicited larger N1 amplitudes than tones, an effect that might be accounted for by differences in the duration of the two sound categories (50 ms for tones; 300 ms for voices): N1 amplitude was found to decrease with a shorter stimulus duration⁵⁴, and to increase linearly with a longer tone duration⁵⁵. However, given the complexity of factors known to modulate the N1 amplitude, other factors may have accounted for this effect (e.g., attention⁵⁴).

Variable	Estimate	SE	t-value	Pr(> z)
Fixed Effects				
Intercept	0.02331	0.01452	1.606	0.114183
LSHS _{Total}	0.00229	0.0005372	3.810	<0.001***
LSHS _{Total} * Self* Tone (vs. External Tone)	0.0006327	0.0003783	1.673	0.095054
LSHS _{Total} * External* Voice (vs. Self-Voice)	-0.001799	0.0003783	-4.756	<0.001***
LSHS _{Total} * Self* Tone (vs. Self-Voice)	-0.001464	0.0003783	-3.870	<0.001***
LSHS* External* Tone (vs. External Voice)	0.0002974	0.0003783	0.786	0.432177
Groups	Name	Variance	SD	
Random Effects				
Subject	(Intercept)	0.001294	0.03597	
Residual		0.001745	0.04177	

Table 2. Linear mixed effects model of pre-stimulus alpha power including the effect of hallucination predisposition. *Notes.* SE = standard error; SD = standard deviation; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Degrees of freedom for Fixed Effects: $df = 480.00$ (except Intercept: $df = 53.44$).

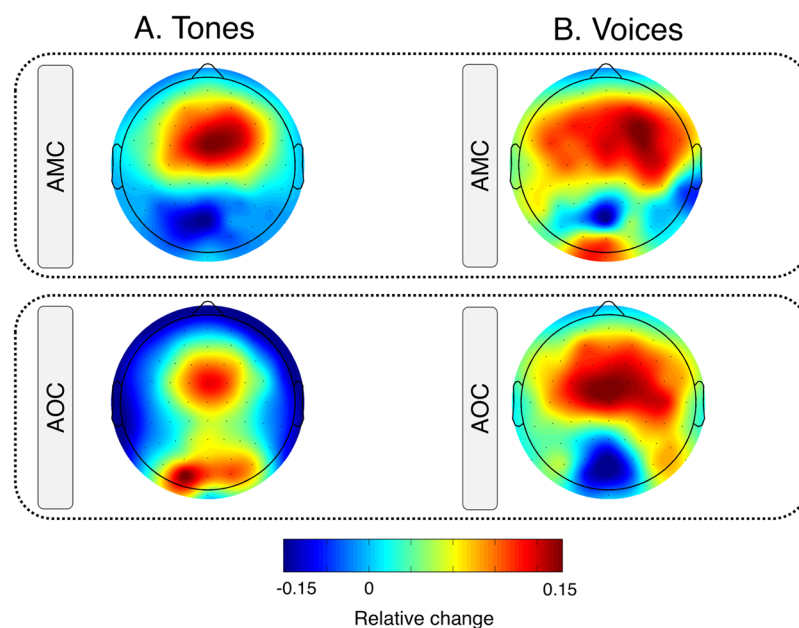


Figure 4. Maps showing the topographical distribution of pre-stimulus alpha power for the different conditions and stimulus types. Values are averaged across the 8–12 Hz frequency band in a time window of 250 ms before stimulus onset.

Whereas the N1 attenuation effect to self-initiated tones was not affected by HP, the response to self-generated voices was. Higher HP resulted in an N1 enhancement rather than suppression: self-generated voices elicited a larger N1 amplitude modulation than externally-generated voices. Typically, an increased N1 response to self-generated auditory stimuli has been related to an increase in prediction error. When sensory feedback mismatches a prediction, the N1 suppression effect is reduced^{3,56,57}. Further, when predictions are less specific, the N1 suppression effect is smaller⁵⁸. Based on this evidence, one could also argue that when voices are less clearly predicted, sensory feedback is not attenuated. Hence, the increased N1 response to self-triggered voices may indicate that the self-voice is less accurately predicted even when participants with higher HP produce sensory feedback to their voices. Specific predictions are even more relevant in the case of more complex auditory stimuli, for which features such as stimulus frequency, onset, and intensity need to be incorporated into the prediction of the self-voice. Thus, a prediction error may arise from the comparison of a less specific/accurate prediction with the available sensory feedback. A less specific prediction may imply altered self-monitoring of speech, which has been consistently associated with AVH^{34,36,37,59–62}.

We cannot rule out the contribution of attentional processes when looking at the enhanced N1 response as a function of increased HP. Increased attention to the to-be-presented voice may have prevented the motor-induced attenuation effect. Cognitive theories of AVH contemplate the role of biased attentional processes in some forms of AVH⁶³. Attention and prediction processes may have opposite effects with regards to the N1 attenuation: whereas prediction reduces the N1, attention results in increased N1 amplitude^{64,65}. An increased N1 response to self-triggered voices due to increased HP may indicate that participants with high HP focus their attention more on self-voice stimuli⁶⁴.

The conscious detection of a self-initiated voice is not affected by HP. The P2 response was also suppressed to both self-initiated tones and voices compared to their passive presentation. The existing evidence is less consistent regarding the effects of predictable sensory feedback on the P2. The pattern we found here is compatible with our previous results on tones^{15,18,19}. The P2 suppression effect may reflect the conscious detection of sensory feedback as self-generated, playing a role in the distinction of self and externally-produced sounds^{15,18,19}. Both groups showed a similar P2 suppression effect for tones and voices. This similarity suggests that the conscious detection of a sound that follows a button-press as self-produced is not affected by HP. This may underlie the preserved sensory experience of a self-generated sound despite reduced sensory suppression. Similar findings (i.e., lack of differences in P2 amplitude relative to controls) were observed in talking or button-press paradigms in schizophrenia¹³ or cerebellar patients^{18,19}.

Increased expectancy for voices in hallucination predisposition. We also observed that HP modulated EEG activity before sound onset. Specifically, HP affected pre-stimulus alpha activity for self-generated voices only: increased HP resulted in increased alpha power. Activity preceding an action (e.g., pressing a button that elicits a sound) has been proposed to reflect the formation of an evidence copy²⁷. Increased pre-stimulus alpha power has been related to prediction effects (occurring in the medial prefrontal cortex²⁸) that allow preparing the auditory cortex for the processing of self-generated sounds^{28,29}. Specifically, alpha power is sensitive to the precision of a prediction regarding an upcoming stimulus. As HP participants did not show the expected sensory attenuation effect to self-triggered voices, it is possible that enhanced predictability to relevant features of the self-voice results in increased synaptic gain that reverses the suppressive effects of the prediction (i.e., the N1 to self-triggered voices fails to be suppressed).

Implications for models of AVH and of a psychosis continuum. Although high HP participants showed a sensory attenuation effect that differs from that previously reported in schizophrenia (lack of a suppression effect)¹¹, the increased N1 response to self-voices is partially consistent with an altered internal forward model in AVH. Whereas HP did not alter the processing of externally generated sounds (no difference between external tone and external voice), it did affect self-voice processing. Further, in both ERP and EEG oscillatory activity, the effects of HP were selective for voices, but not tones. A voice-specific rather than generalized, altered sensory prediction may explain why individuals who experience AVH hallucinated voices display the most common abnormal perceptual experience.

It is worth noting that even though this was not explicitly reported in talking paradigms (i.e., in which voice stimuli were tested) in schizophrenia patients with AVH, the inspection of grand average waveforms suggests that HP affected the processing of self-generated vocal sounds more (N1 amplitude was more negative in the talking condition in patients with hallucinations compared to controls) than the processing of externally generated voices (no group differences)¹¹. It is possible that along the psychosis continuum, altered sensory feedback that can be predicted as a function of one's own action, occurs first for auditory stimuli with increased social relevance (e.g., voices) and in clinical stages of the continuum it generalizes to other types of sounds (e.g., simple tones). This hypothesis is admittedly speculative and needs to be tested in future studies with larger samples.

Together, the current findings suggest that sensory feedback to self-voice is altered in people with an AVH predisposition. Specific alterations in the processing of self-generated vocal sounds may thus establish a core feature on the psychosis continuum.

Data Availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Wolpert, D. M. & Kawato, M. Multiple paired forward and inverse models for motor control. *Neural Networks* **11**, 1317–1329 (1998).
2. Wolpert, D. M., Miall, R. C. & Kawato, M. Internal models in the cerebellum. *Trends Cogn. Sci.* **2**, 338–347 (1998).
3. Behroozmand, R., Karvelis, L., Liu, H. & Larson, C. R. Vocalization-induced enhancement of the auditory cortex responsiveness during voice F0 feedback perturbation. *Clin. Neurophysiol.* **120**, 1303–1312 (2009).
4. Curio, G., Neuloh, G., Numminen, J., Jousmäki, V. & Hari, R. Speaking modifies voice-evoked activity in the human auditory cortex. *Hum. Brain Mapp.* **9**, 183–191 (2000).
5. Ford, J. M. *et al.* Neurophysiological evidence of corollary discharge dysfunction in schizophrenia. *Am. J. Psychiatry* **158**, 2069–2071 (2001).
6. Ford, J. M. & Mathalon, D. H. Corollary discharge dysfunction in schizophrenia: Can it explain auditory hallucinations? *Int. J. Psychophysiol.* **58**, 179–189 (2005).
7. Ford, J. M. & Mathalon, D. H. Electrophysiological evidence of corollary discharge dysfunction in schizophrenia during talking and thinking. *J. Psychiatr. Res.* **38**, 37–46 (2004).
8. Heinks-Maldonado, T. H., Nagarajan, S. S. & Houde, J. F. Magnetoencephalographic evidence for a precise forward model in speech production. *Neuroreport* **17**, 1375–9 (2006).
9. Ford, J. M., Roach, B. J. & Mathalon, D. H. Assessing corollary discharge in humans using noninvasive neurophysiological methods. *Nat. Protoc.* **5**, 1160–1168 (2010).
10. Ford, J. M., Roach, B. J., Faustman, W. O. & Mathalon, D. H. Synch before you speak: Auditory hallucinations in schizophrenia. *Am. J. Psychiatry* **164**, 458–466 (2007).
11. Heinks-Maldonado, T. H. *et al.* Relationship of imprecise corollary discharge in schizophrenia to auditory hallucinations. *Arch. Gen. Psychiatry* **64**, 286–296 (2007).
12. Baess, P., Widmann, A., Roye, A., Schröger, E. & Jacobsen, T. Attenuated human auditory middle latency response and evoked 40-Hz response to self-initiated sounds. *Eur. J. Neurosci.* **29**, 1514–1521 (2009).
13. Ford, J. M., Palzes, V. A., Roach, B. J. & Mathalon, D. H. Did I do that? Abnormal predictive processes in schizophrenia when button pressing to deliver a tone. *Schizophr. Bull.* **40**, 804–812 (2014).
14. Hughes, G., Desantis, A. & Waszak, F. Attenuation of auditory N1 results from identity-specific action-effect prediction. *Eur. J. Neurosci.* **37**, 1152–1158 (2013).

15. Knolle, F., Schröger, E. & Kotz, S. A. Prediction errors in self- and externally-generated deviants. *Biol. Psychol.* **92**, 410–416 (2013).
16. Crapse, T. B. & Sommer, M. A. Corollary discharge across the animal kingdom. *Nat. Rev. Neurosci.* **9**, 587–600 (2008).
17. Feinberg, I. Efference Copy and Corollary Discharge: Implications for Thinking and Its Disorders*. *Schizophr. Bull.* **4**, 636–640 (1978).
18. Knolle, F., Schröger, E. & Kotz, S. A. Cerebellar contribution to the prediction of self-initiated sounds. *Cortex* **49**, 2449–2461 (2013).
19. Knolle, F., Schröger, E., Baess, P. & Kotz, S. A. The Cerebellum Generates Motor-to-Auditory Predictions: ERP Lesion Evidence. *J. Cogn. Neurosci.* **24**, 698–706 (2012).
20. Godey, B., Schwartz, D., De Graaf, J. B., Chauvel, P. & Liégeois-Chauvel, C. Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: A comparison of data in the same patients. *Clin. Neurophysiol.* **112**, 1850–1859 (2001).
21. Näätänen, R. & Michie, P. T. Early selective-attention effects on the evoked potential: A critical review and reinterpretation. *Biol. Psychol.* **8**, 81–136 (1979).
22. Zouridakis, G., Simos, P. G. & Papanicolaou, A. C. Multiple bilaterally asymmetric cortical sources account for the auditory N1m component. *Brain Topogr.* **10**, 183–189 (1998).
23. Ford, J. M. & Mathalon, D. H. Anticipating the future: Automatic prediction failures in schizophrenia. *Int. J. Psychophysiol.* **83**, 232–239 (2012).
24. Picard, F. & Friston, K. Predictions, perception, and a sense of self. *Neurology* **83**, 1112–8 (2014).
25. Wolpert, D. M. & Flanagan, J. R. Primer: Motor Prediction. *Curr. Biol.* **11**, 729–732 (2001).
26. Chen, Z., Chen, X., Liu, P., Huang, D. & Liu, H. Effect of temporal predictability on the neural processing of self-triggered auditory stimulation during vocalization. *BMC Neurosci.* **13**, 55 (2012).
27. Wang, J. *et al.* Action planning and predictive coding when speaking. *Neuroimage* **91**, 91–98 (2014).
28. Müller, N., Leske, S., Hartmann, T., Szebényi, S. & Weisz, N. Listen to yourself: The medial prefrontal cortex modulates auditory alpha power during speech preparation. *Cereb. Cortex* **25**, 4029–4037 (2015).
29. Cao, L., Thut, G. & Gross, J. The role of brain oscillations in predicting self-generated sounds. *Neuroimage* **147**, 895–903 (2017).
30. Stenner, M.-P., Bauer, M., Haggard, P., Heinze, H.-J. & Dolan, R. Enhanced Alpha-oscillations in Visual Cortex during Anticipation of Self-generated Visual Stimulation. *J. Cogn. Neurosci.* **26**, 2540–2551 (2014).
31. Verdoux, H. & Van, O. J. Psychotic symptoms in non-clinical populations and the continuum of psychosis. *Schizophr. Res.* **54**, 59–65 (2002).
32. Beavan, V., Read, J. & Cartwright, C. The prevalence of voice-hearers in the general population: A literature review. *J. Ment. Heal.* **20**, 281–292 (2011).
33. Linscott, R. J. & van Os, J. An updated and conservative systematic review and metaanalysis 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**, 1133–1149 (2013).
34. Brébion, G. *et al.* Impaired self-monitoring of inner speech in schizophrenia patients with verbal hallucinations and in non-clinical individuals prone to hallucinations. *Front. Psychol.* **7**, 1381 (2016).
35. Barkus, E., Stirling, J., Hopkins, R., McKie, S. & Lewis, S. Cognitive and neural processes in non-clinical auditory hallucinations. *Br. J. Psychiatry* **51**, s76–81 (2007).
36. Allen, P. *et al.* Misattribution of external speech in patients with hallucinations and delusions. *Schizophr. Res.* **69**, 277–287 (2004).
37. Pinheiro, A. P. *et al.* Emotional self-other voice processing in schizophrenia and its relationship with hallucinations: ERP evidence. *Psychophysiology* **54**, 1252–1265 (2017).
38. Ford, J. M., Roach, B. J., Faustman, W. O. & Mathalon, D. H. Out-of-Synch and Out-of-Sorts: Dysfunction of Motor-Sensory Communication in Schizophrenia. *Biol. Psychiatry* **63**, 736–743 (2008).
39. Moseley, P., Smailes, D., Ellison, A. & Fernyhough, C. The effect of auditory verbal imagery on signal detection in hallucination-prone individuals. *Cognition* **146**, 206–216 (2016).
40. Alderson-Day, B. *et al.* Distinct processing of ambiguous speech in people with non-clinical auditory verbal hallucinations. *Brain* **9**, 2475–2489 (2017).
41. Conde, T., Gonçalves, O. F. & Pinheiro, A. P. A Cognitive Neuroscience view of voice processing abnormalities in schizophrenia: a window into auditory verbal hallucinations? *Harv. Rev. Psychiatry* **24**, 148–163 (2016).
42. Castiájo, P. & Pinheiro, A. P. On 'hearing' voices and 'seeing' things: Probing hallucination predisposition in a Portuguese nonclinical sample with the Launay-Slade Hallucination Scale-revised. *Front. Psychol.* **8**, 1138 (2017).
43. Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39**, 175–191 (2007).
44. Oldfield, R. C. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* **9**, 97–113 (1971).
45. Pinheiro, A. P., Rezaii, N., Rauber, A. & Niznikiewicz, M. 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 (2016).
46. Bates, D., Maechler, M., Bolker, B. & Walker, S. lme4: linear mixed-effects models using Eigen and S4. *J. Stat. Softw.* **67**, 1–48 (2015).
47. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest: Tests in linear mixed effects models. *R package version 2*, 0–33 (2016).
48. Baayen, R. H., Davidson, D. J. & Bates, D. M. Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* **59**, 390–412 (2008).
49. Bates, D., Mächler, M., Bolker, B. M. & Walker, S. C. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
50. Näätänen, R. & Picton, T. The N1 Wave of the Human Electric and Magnetic Response to Sound: A Review and an Analysis of the Component Structure. *Psychophysiology* **24**, 375–425 (1987).
51. Woods, D. L. The component structure of the N1 human auditory evoked potential. *Electroencephalogr. Clin. Neurophysiol.* **44**, 102–109 (1995).
52. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Automat. Contr.* **19**, 716–723 (1974).
53. Whitford, T. J. *et al.* Electrophysiological and diffusion tensor imaging evidence of delayed corollary discharges in patients with schizophrenia. *Psychol. Med.* **41**, 959–969 (2011).
54. Rosburg, T., Boutros, N. N. & Ford, J. M. Reduced auditory evoked potential component N100 in schizophrenia - A critical review. *Psychiatry Res.* **161**, 259–274 (2008).
55. Ostroff, J. M., McDonald, K. L., Schneider, B. A. & Alain, C. Aging and the processing of sound duration in human auditory cortex. *Hear. Res.* **181**, 1–7 (2003).
56. Behroozmand, R., Liu, H. & Larson, C. R. Time-dependent Neural Processing of the Auditory Feedback during Voice Pitch Error Detection. *J. Cogn. Neurosci.* **23**, 1–13 (2007).
57. Behroozmand, R. & Larson, C. R. Error-dependent modulation of speech-induced auditory suppression for pitch-shifted voice feedback. *BMC Neurosci.* **12**, 54 (2011).
58. Baess, P., Jacobsen, T. & Schröger, E. Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: Evidence for internal forward models with dynamic stimulation. *Int. J. Psychophysiol.* **70**, 137–143 (2008).
59. Allen, P., Freeman, D., Johns, L. & McGuire, P. Misattribution of self-generated speech in relation to hallucinatory proneness and delusional ideation in healthy volunteers. *Schizophr. Res.* **84**, 281–288 (2006).
60. Allen, P. *et al.* Neural correlates of the misattribution of speech in schizophrenia. *Br. J. Psychiatry* **190**, 162–169 (2007).
61. Johns, L. C. *et al.* Impaired verbal self-monitoring in individuals at high risk of psychosis. *Psychol. Med.* **40**, 1433–1442 (2010).

62. Johns, L. C. *et al.* Verbal self-monitoring and auditory verbal hallucinations in patients with schizophrenia. *Psychol. Med.* **31**, 705–715 (2001).
63. Hugdahl, K. *et al.* Auditory hallucinations in schizophrenia: the role of cognitive, brain structural and genetic disturbances in the left temporal lobe. *Front. Hum. Neurosci.* **1**, 6 (2008).
64. Lange, K. The ups and downs of temporal orienting: a review of auditory temporal orienting studies and a model associating the heterogeneous findings on the auditory N1 with opposite effects of attention and prediction. *Front. Hum. Neurosci.* **7**, 263 (2013).
65. Schroger, E., Kotz, S. A. & SanMiguel, I. Bridging prediction and attention in current research on perception and action. *Brain Res.* **1626**, 1–13 (2015).

Acknowledgements

The Authors gratefully acknowledge all the participants who collaborated in the study, and particularly Dr. Franziska Knolle for feedback on stimulus generation, Carla Barros for help with scripts for EEG time-frequency analysis, and Dr. Célia Moreira for her advice on mixed linear models. This work was supported by the Portuguese Science National Foundation (FCT; grant numbers PTDC/PSI-PCL/116626/2010, IF/00334/2012, PTDC/MHC-PCN/0101/2014) awarded to APP.

Author Contributions

A.P.P., M.S. and S.A.K. developed the study concept and design. A.P.P. collected the data. The three authors collaborated in data analysis and in writing the first draft of the manuscript. They equally contributed to and have approved the final version of the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-32614-9>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018