Dissociable effects of agency and ownership on speech perception

Margarida Marques[1], Maria Amorim[1], Sonja A. Kotz[2*], & Ana P. Pinheiro[1*]

[1]Faculty of Psychology, University of Lisbon, Lisbon, Portugal

[2]Department of Neuropsychology and Psychopharmacology, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, Netherlands

* Contributed equally

Corresponding author:

Ana P. Pinheiro

Faculty of Psychology, University of Lisbon,

Alameda da Universidade, 1649-013 Lisboa, Portugal

e-mail: appinheiro@psicologia.ulisboa.pt

Highlights

- A button-press EEG task with self-/other-voice feedback dissociated agency and ownership.
- N1 amplitude indexed voice ownership; P2 reflected agency attribution.
- Source localization revealed distinct neural networks for agency and ownership.
- Results support a hybrid model integrating sensorimotor predictions and voice representations.
- Findings may guide EEG-based tools for detecting self-related dysfunctions.

Abstract

Speech production relies on a forward model that predicts the sensory consequences of one's vocal actions. When auditory feedback matches these predictions, event-related potentials (ERPs) of the electroencephalogram (EEG) are attenuated, supporting two partly overlapping experiences: agency (the sense of causing the action) and ownership (the sense that the voice belongs to oneself). Whether early sensory attenuation effects dissociate these experiences is unknown.

We recorded EEG from 50 participants during a voice playback task. Agency was manipulated by contrasting self-chosen (self-generated) and unchosen (other-generated) words; ownership contrasted the participant's own voice and the experimenter's voice. Sensory attenuation of the N1 and P2 ERPs was measured, and their cortical generators were source-localized with eLORETA.

N1 amplitude was selectively modulated by ownership, whereas the P2 indexed agency, suggesting that ownership precedes and subsequently impacts agency. Source localization revealed partially dissociable networks: ownership engaged a fronto-temporal network, including the insula, whereas agency recruited inferior parietal and somatosensory regions.

These findings reveal a temporal and spatial dissociation between voice agency and ownership. They further suggest that the forward model integrates both sensorimotor predictions and memory-based representations to support a coherent sense of self in vocal communication.

Keywords: agency, ownership, voice, forward model.

1. Introduction

The voice is a powerful social signal and a unique marker of individuality. Like a fingerprint, each voice is acoustically distinct, encoding physical, psychological, and social information about the speaker (Lavan, 2023; Latinus & Belin, 2011). When we speak, two distinct aspects of the self are at play: initiating vocal output (agency) and recognizing the voice as one's own (ownership) (Di Plinio et al., 2020; Ohata et al., 2022; Synofzik et al., 2008). When we hear ourselves speak, how do we distinguish between causing a vocal sound and identifying it as self-produced? And how are these signals integrated to support efficient self-other distinction?

1.1. Voice Agency: I am the one who speaks

Self-voice perception relies on motor-based predictions during speech production. A forward model is thought to facilitate the distinction between self-generated and externally generated sensory input (Schröger et al., 2015; Wolpert, 1997). In speech, an efference copy of the motor command is relayed from motor planning regions to the auditory cortex via the cerebellum, allowing the brain to anticipate the sensory consequences of speaking (Miall & Wolpert, 1996; Pinheiro et al., 2020).

If auditory feedback matches the predicted outcome, the sensory experience is perceived as self-generated (Gentsch & Schütz-Bosbach, 2011; Picard & Friston, 2014; Weiss et al., 2011), often leading to sensory attenuation: self-produced sounds are perceived as less intense than identical external sounds. At a neurophysiological level, this is indexed by reduced amplitude of the N1 event-related potential (ERP) around 100 ms post-stimulus (Knolle et al., 2019; Pinheiro et al., 2020).

The P2 component has also been linked to agency, albeit less consistently (Beño-Ruiz-de-la-Sierra et al., 2023; Poonian et al., 2015). While the N1 and P2 are often treated as a functional complex (e.g., Wagner et al., 2017), recent work highlights their distinct roles (Egan et al., 2023; Knolle et al., 2012; Seidel et al., 2021). Unlike the N1, which is typically modulated regardless of conscious self-attribution, the P2 may reflect higher-order aspects of self-causation, including the conscious experience of agency (Seidel et al., 2021; Timm et al., 2016).

1.2. Voice Ownership: The voice is mine

Ownership arises when the voice is recognized as belonging to the self. In speech production, a match between predicted and actual auditory feedback can foster the sense of ownership, as reflected in N1 attenuation (Behroozmand et al., 2009; Heinks-Maldonado et al., 2007). However, ownership can also arise independently of motor signals. Passive listening studies show that individuals can identify their own voice in the absence of an efference copy (Pinheiro et al., 2016), suggesting that voice ownership also draws on long-term identity representations (Andics et al., 2013; Niziolek et al., 2013). Consistently, the N1 is modulated by voice identity even during passive listening, often showing increased amplitude to one's own voice (Pinheiro et al., 2023).

1.3. Integration and Dissociation of Agency and Ownership

Although agency and ownership typically co-occur in speech, they are dissociable both conceptually and neurally (Hubl et al., 2014; Kalckert & Ehrsson, 2012; Synofzik et al., 2008). A key functional magnetic resonance imaging (fMRI) study by Di Plinio and colleagues (2020) orthogonally manipulated voice feedback along two dimensions: whether feedback matched the participant's choice (agency), and whether it was spoken in their own voice (ownership). Agency and ownership engaged distinct neural networks: ownership activated temporal auditory and insular cortices, whereas agency engaged a broader sensorimotor network, including the angular gyrus, intraparietal sulcus, inferior frontal gyrus, supplementary motor area, precuneus, and cerebellum.

However, the temporal dynamics of this dissociation remain unclear. Time-resolved EEG markers such as the N1 and P2 can help determine when and how these processes diverge during speech, but findings are mixed. Some studies link N1 suppression to agency (Pinheiro et al., 2020), whereas others identify the P2 as a more robust marker (Timm et al., 2016). Similarly, while the N1 has been associated with ownership (Behroozmand & Larson, 2011; Pinheiro et al., 2023), its specificity remains debated (Beño-Ruiz-de-la-Sierra et al., 2023).

Clarifying these temporal dynamics is particularly relevant in clinical contexts. Altered responses to the self-voice are observed in individuals prone to auditory hallucinations (Allen et al., 2006; Johns et al., 2010; Pinheiro et al., 2019a) and associated with disrupted agency attribution (Bühler et al., 2016; Beño-Ruiz-de-la-Sierra et al., 2024). Dissociating voice agency and ownership may enhance our understanding of self-disturbances across conditions such as schizophrenia, dissociative disorders, and autism (Uhlmann et al., 2021; Zhao et al., 2013).

1.4. The Present Study

Previous attempts to dissociate agency and ownership in speech production have been limited by key methodological constraints. Many studies isolate either the temporal or spatial dimensions of these processes, overlooking how their interaction may jointly contribute to coherent self-related processing. Moreover, agency is often examined in overly simplified contexts, with limited attention to interpersonal or ecologically valid scenarios. This has led to a narrow focus on low-level motor control, neglecting the higher-order dimensions of agency attribution.

This EEG study addressed these limitations using EEG combined with exact low-resolution brain electromagnetic tomography (eLORETA) to disentangle the spatiotemporal neural dynamics of voice agency and ownership. Building on a previous fMRI paradigm (Di Plinio et al., 2020), a more interactive setup was created, where participants believed they were interacting with the experimenter. Agency was manipulated via outcome contingency: auditory feedback either matched the participant's word choice (self-generated) or was attributed to the experimenter (other-generated). Critically, motor preparation and predictability were held constant across conditions (Desantis et al., 2012). Voice ownership, in turn, was manipulated via speaker identity: feedback was delivered in either the participant's own voice or the experimenter's voice.

By analyzing N1 and P2 responses separately, we aimed to track the unfolding of agency and ownership with high temporal precision. We hypothesized that both agency and ownership would independently modulate the N1, indexing early self-other discrimination. The P2 was expected to index the subjective experience of agency. At the source level, we predicted that ownership would engage temporal auditory cortex and insula, while agency would recruit frontoparietal and motor regions (Di Plinio et al., 2020).

2. Method

We provide a detailed account of how our sample size was determined, clearly specify all inclusion criteria, and describe all experimental manipulations and measures used in the study.

2.1. Participants

Fifty-four young adults (25 males; mean age = 22.22 years, *SD* = 3.60, age range 18-32) who were native speakers of European Portuguese participated in the experiment. The inclusion/exclusion criteria were established prior to data analysis. Participants were required to meet the following inclusion criteria: (1) age between 18 and 32 years; (2) normal hearing;

(3) normal or corrected-to-normal vision; and (4) no reported history of neurological or psychiatric disorders. Four were excluded due to EEG artifacts, resulting in a final sample of 50 participants (25 males; mean age = 22.40 years, *SD* = 3.65; age range 18-32).

The target sample size was not determined through a priori power analysis, as reliable effect size estimates could not be derived from previous research due to the scarcity of directly comparable studies. Instead, recruitment continued for the full duration of the allotted data collection period, with the aim of obtaining a sample size consistent with prior ERP investigations of self-voice perception (e.g., Iannotti et al., 2022; Pinheiro et al., 2016).

The study was approved by the Ethics Committee of the Faculty of Psychology, University of Lisbon, and conducted in accordance with the Declaration of Helsinki. All participants provided written informed consent and received vouchers and/or course credit.

2.2. Stimuli

We selected 68 semantically neutral disyllabic words (e.g., ‘mesa’ / ‘table’) from the European Portuguese adaptation of the Affective Norms for English Words (ANEW; Soares et al., 2012). Words with extreme valence ratings (below 4 or above 6), unusual length for disyllables (fewer than 4 or more than 6 letters), or unsuitable grammatical class were excluded (e.g., verbs). From the resulting 90 words, 34 matched pairs were created based on syllable structure and letter count, yielding 68 words for the experiment.

2.3. Procedure

Participants completed three sessions in the following order: (1) voice recording, (2) online behavioral task, (3) EEG session. Prior to participation, sociodemographic information was collected via an online questionnaire. No part of the study procedures was pre-registered prior to the research being conducted.

2.3.1. Voice Recording

Voice recordings took place in a sound-attenuating chamber using a Roland R-16 digital recorder (44,100 kHz sampling rate; 16-bit quantization). Participants read each of the 68 words aloud. The words appeared in the center of a computer screen sequentially and participants were instructed to repeat each word five times after listening to a standardized model utterance recorded with neutral prosody and constant loudness. This session took around 40 minutes.

Background noise was removed from the recordings using Audacity software. Words were then segmented using Praat software. Their intensity was normalized (65 dB) using a Praat script. Following the same procedure, the words were also recorded by two experimenters (one male, age 27; one female, age 25) to generate other-voice tokens, matched by the participant's gender.

2.3.2. Online Behavioral Session

To ensure reliable self-other voice discrimination, participants completed a forced-choice task using voice morphs created with TANDEM-STRAIGHT (Kawahara & Morise, 2011) in MATLAB. Four continua were generated (11 steps, 0-100% self). Participants judged whether each stimulus sounded "more mine" or "more other". Only responses to unmorphed voices were analyzed here. High accuracy (self-voice: *M* = 93.21%; other-voice: 91.67%) confirmed efficient self-other voice distinction.

2.3.3. EEG Session

The EEG session occurred at least one week after the voice recording session. Participants sat 80 cm from a computer screen in a soundproof chamber. To minimize movement artifacts, we used a button-press task that produces comparable findings to talk-listen paradigms (Knolle et al., 2019; Pinheiro et al., 2020).

On each trial, a word pair (randomly selected from the 34 pairs) was displayed in white text. After 3000 ms, the words turned red, cueing participants to select one silently and indicate their choice via keyboard (keys ‘F’/’K’, left/right positions counterbalanced). If no response was made within 1500 ms, the trial proceeded automatically without auditory feedback.

Auditory feedback (a spoken word) was presented 1000 ms after the button-press, alongside a central fixation cross. This delay minimized motor contamination of sound-locked EEG activity (Poonian et al., 2015). Voice feedback could either be the chosen word or an unchosen word, uttered in one’s own or in the experimenter’s voice, yielding four experimental conditions: SS: self-chosen word, self-voice; SO: self-chosen word, other-voice; OS: other-chosen word, self-voice; OO: other-chosen word; other-voice. The paradigm is illustrated in Figure 1.

Before starting the EEG task, participants completed a training phase in which a shortened version of the task was administered, and they received consistent feedback matching their choices. In the main task, however, they were told that the experimenter could override their choices remotely, enhancing believability of the agency manipulation.

Each word pair was presented four times, totaling 136 randomized trials. Stimuli occurred with equal probability across conditions. The task was implemented in Presentation (Version 23; Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com/).

2.4. EEG Data Acquisition and Preprocessing

EEG was continuously recorded with a 64-channel BioSemi ActiveTwo system (sampling rate: 512 Hz). Using Brain Vision Analyzer 2 (www.brainproducts. com/), data were re-referenced offline to the average of the mastoids and filtered with a 0.1–30 Hz bandpass filter (zero phase shift Butterworth filters; order 2). EEG epochs were time-locked to auditory feedback onset, including a 200 ms pre-stimulus baseline and extended to 1000 ms post-stimulus, and baseline-corrected using the pre-stimulus interval. Vertical and horizontal eye movement artifacts were eliminated using the method developed by Gratton, Coles, and Donchin (1983). EEG epochs exceeding ±100 µV were rejected. At least 70% of segments per condition were retained per participant. There were no significant differences in the number of trials across conditions ($p > .05$). Lastly, grand average waveforms were created for each condition (Figure 2).

No part of the study analyses was pre-registered prior to the research being conducted.

2.5. Statistical Analyses of ERP Data

Time windows for ERP analysis were identified using the collapsed-localizer method (Luck & Gaspelin, 2017). Mean amplitudes were extracted for the N1 (120-200 ms) and P2 (210-320 ms) components from frontal (F3, Fz, F4), frontocentral (FC3, FCz, FC4), and central (C3, Cz, C4) electrodes, based on their established topographies in motor-auditory paradigms (Pinheiro et al., 2019b).

Linear mixed-effects models (LMMs) assessed the effects of agency and ownership, with random intercepts for participants, using the lmer4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017) R packages (R4.4.1. GUI 1.81). The anova function from lmerTest (Kuznetsova et al., 2017) provided *p*-values for each factor, calculated from the *F*-statistics. The Satterthwaite approximation was applied to the REML-fitted models (Luke, 2017).

Compared to the traditional repeated-measures ANOVA analysis, mixed-effects modeling can account for ERP amplitude variance stemming from random factors, such as participant-level variability (Baayen et al., 2008). Given that it can also produce more efficient estimates and powerful tests, mixed-effects modelling has been increasingly recommended for the analysis of repeated-measures experiments (Bagiella et al., 2000; Boisgontier & Cheval, 2016).

2.6. Source Localization Analyses (eLORETA)

eLORETA (Pascual-Marqui, 2007; http://www.uzh.ch/keyinst/loreta.htm) was used to estimate cortical sources of ownership- and agency-related EEG effects. Analyses covered the 1000 ms post-stimulus interval (Pinal et al., 2014).

Operating within a standardized head model based on the Montreal Neurological Institute (MNI152) template, eLORETA partitions the cortical gray matter into 6239 voxels (5 mm spatial resolution) and calculates current density estimates at each voxel (i.e., the electric activity at each voxel, which serves as a proxy for neuronal activity) through a transformation matrix (Pascual-Marqui, 2007). More specifically, eLORETA computes a three-dimensional distribution of intracortical current density that minimizes localization error while maximizing spatial smoothness. This computation is based on the physiologically plausible assumption that neighboring cortical regions tend to exhibit correlated activation (Pascual-Marqui, 2002).

Statistical comparisons between conditions were performed using two-tailed, non-parametric paired-samples t-tests on a timepoint-by-timepoint basis for all electrodes (Esslen et al., 2008). Statistical non-Parametric Mapping (SnPM) was used with 5000 random permutations to control for multiple comparisons (Nichols & Holmes, 2002). Timeframes showing significant differences ($p < .05$) that overlapped with the latencies of the ERP components of interest were subsequently used for source-level statistical testing.

eLORETA transformation matrix was applied to scalp-recorded EEG data to generate voxel-wise estimated current density maps for each participant per condition. Using these maps and within the previously identified significant time intervals, the software computed voxel-wise paired-sample statistics based on the log of F-ratio (LnF) (Esslen et al., 2008; Takahashi et al., 2013).

Statistical inference was then carried out using SnPM with 5000 randomizations (Nichols & Holmes, 2002). This procedure yielded maps of LnF values, identifying cortical regions that exhibit differences in estimated current density between experimental conditions based on a threshold value corresponding to a p-value of .05.

3. Results

3.1. ERP data

3.1.1. N1

N1 amplitude was significantly modulated by ownership ($F(1, 1750) = 19.504$, $p < .001$). Specifically, N1 amplitude was larger (i.e., more negative) for words spoken in the self-voice than in the other-voice ($\beta = .374$, SE = .118, $t = 3.161$, $p = .002$, 95% CI [.142, .606]). Agency did not modulate the N1 ($F(1, 1750) = .002$, $p = .969$): N1 amplitude to self-chosen words was not significantly different from other-chosen words. In addition, the effect of ownership did not interact with agency ($F(1, 1750) = .003$, $p = .957$).

3.1.2. P2

The results confirmed an agency effect in the P2 time window ($F(1, 1750) = 36.925$, $p < .001$): self-chosen words elicited an enhanced P2 response relative to other-chosen words ($\beta = -1.021$, SE = .172, $t = -5.946$, $p < .001$, 95% CI [-1.358, -.685]). P2 amplitude was not affected by ownership ($F(1, 1750) = .011$, $p = .918$). However, a significant interaction between agency and ownership emerged ($F(1, 1750) = 5.795$, $p = .016$): the agency effect on the P2 was larger for self-voice stimuli than for other-voice stimuli ($\beta = .585$, SE = .243, $t = 2.407$, $p = .016$, 95% CI [.109, 1.061]). This interaction indicates that the agency effect was modulated by the degree of self-relevance signaled by voice identity.

3.2. Source Localization

3.2.1. Self-voice vs. Other-voice Contrast

Between 121-139 ms post-stimulus, increased current source density estimates were observed in response to the self-voice compared to the other-voice in left temporal regions ($p < .05$). Specifically, enhanced activation was localized to the middle temporal gyrus (MTG; x = -50, y = 0, z = 15), superior temporal gyrus (STG; x = -50, y = 5, z = 15), and inferior temporal gyrus (ITG; x = -60, y = -5, z = -20). Greater estimated activation was also observed in a left temporal sub-gyral area (x = -40, y = 0, z = -10), as well as in the left insula (x = -40, y = 0, z = -5) and the left inferior frontal gyrus (IFG; x = -40, y = 15, z = -15) (Figure 3; see Table 1 for a summary of findings).

3.2.2. Self-chosen vs. Other-chosen Word Contrast

At latencies ranging from 307 to 320 ms, self-generated words elicited stronger source activity in the left inferior parietal lobule (IPL; x = -50, y = -30, z = 45), the left postcentral gyrus (x = -55, y = -25, z = 50), and the left angular gyrus (x = -30, y = -60, z = 25) (Figure 4; see Table 2 for a summary of findings).

3.2.3. Agency x Ownership Interaction Contrast

To identify the cortical sources underlying the observed P2 interaction between agency and ownership, the contrast (SS – OS) versus (SO – OO) was computed and a voxel-wise paired-sample test based on the *t*-statistic (LnF values are not available for interaction testing in eLORETA) was performed. Significant differences emerged ($p < .05$) between 270 and 275 msec, with stronger estimated activation for the agency effect during self-voice trials in the left anterior cingulate gyrus (ACG; x = -10, y = 45, z = 5) and left medial frontal gyrus (MFG; x =

-10, y = 55, z = 5) (Figure 5; see Table 3 for a summary of findings).

4. Discussion

The present study aimed to dissociate the neural correlates of agency and ownership during speech feedback monitoring using a novel EEG paradigm. The findings reveal a functional and temporal dissociation between these two core components of the self. Specifically, the N1 component was selectively modulated by voice ownership (identity), whereas the P2 component was more sensitive to agency (outcome contingency). These findings were corroborated by source localization analyses, which further identified distinct neural networks for each dimension. Importantly, an interaction between agency and ownership emerged in the P2 time window, suggesting that while these processes can be dissociated, they also converge during self-monitoring in speech.

4.1. The N1 as a Sensitive Marker of Voice Ownership

The enhanced N1 response for the self-voice relative to the other-voice suggests early sensitivity to speaker identity. This finding aligns with previous passive listening studies (Pinheiro et al., 2016, 2023), yet contrasts with the typical N1 suppression observed during overt speech (Behroozmand et al., 2009; Heinks-Maldonado et al., 2007). One possible explanation is attentional prioritization: the N1 is reliably amplified for self-relevant stimuli (Liu et al., 2016; Sui et al., 2023), and the self-voice, being identity-relevant, may automatically capture more attention. This attentional effect could have been further amplified by a perceptual mismatch between the natural and digitally recorded self-voice (e.g., Orepic et al., 2023).

An alternative – but complementary – explanation is that the N1 enhancement reflects early identity recognition via long-term voice templates. Self-voice recognition has been shown to engage a comparison between incoming auditory input and stored internal representations of one's own voice, even in the absence of motor-related predictions (Iannotti et al., 2022; Niziolek et al., 2013; Orepic et al., 2023). This suggests that voice ownership relies not only on sensorimotor predictions but also on memory-based identity cues.

Source localization supported this interpretation, revealing activation in a distributed frontotemporal-insular network, including the middle, superior, and inferior temporal gyri, inferior frontal gyrus, and insula. These regions have been implicated in auditory identity processing, self-monitoring, and multisensory integration (Fu et al., 2006; Johnson et al., 2021; Di Plinio et al., 2020), reinforcing the idea that voice ownership engages a broad network dedicated to distinguishing self from other in the auditory modality.

4.2. The P2 as a Sensitive Marker of Agency

The P2 component was primarily modulated by agency, with larger amplitudes for self-chosen words compared to other-chosen words. This finding supports the view that the P2 reflects the higher-order evaluation of action outcomes (Han et al., 2021; Seidel et al., 2021; Timm et al., 2016). Although the functional significance of the P2 is less clearly defined than that of the N1, it has been linked to the subjective experience of agency, showing greater amplitude suppression when participants perceive themselves as the cause of auditory events (Timm et al., 2016), but also amplitude enhancement in active *vs.* passive conditions (Han et al., 2021).

Given the association of the P2 with attention, stimulus evaluation, and working memory (Crowley & Colrain, 2004; Duzcu et al., 2019; Knolle et al., 2012, 2013), the enhanced response to self-chosen words may reflect increased engagement of evaluative mechanisms that monitor whether sensory outcomes align with internal goals or intentions – core to agency attribution. The act of selecting a word likely reinforced the sense of self-causation and strengthened agency-related processing. This is consistent with models of agency that emphasize postdictive, inferential processes alongside predictive motor signals (Synofzik et al., 2008).

At the source level, agency-related effects were localized to a parietal-sensory network, including the inferior parietal lobule, postcentral gyrus, and angular gyrus – regions consistently associated with action-outcome monitoring and agency attribution (Chambon et al., 2015; DiPlinio et al., 2020; Haggard, 2017; Seghezzi et al., 2019). The absence of frontal or cerebellar sources, typically involved in motor preparation and prediction (Di Plinio et al., 2020; Welniarz et al., 2021), may reflect the study's delayed feedback design, which minimized motor-related contributions to the auditory-locked EEG signal. This pattern may also reflect the spatial resolution limitations of EEG-based source estimation (Saha et al., 2015).

4.3. Integration of Agency and Ownership in the P2

In addition to their dissociation, agency and ownership also interacted during the P2 time window. The amplitude difference between self- and other-chosen words was greater when the feedback was delivered in one's own voice, suggesting that voice identity (ownership) modulates agency attribution. This finding aligns with prior behavioral evidence showing that self-related identity cues influence perceived control over sensory outcomes (Ohata et al., 2022). These results support a model in which ownership cues – associated with identity recognition – exert top-down influences on agency judgments by shaping the perceived relevance of sensory feedback.

Source localization revealed that this interaction engaged the anterior cingulate gyrus and medial frontal gyrus, key regions involved in self-monitoring and self-awareness (Allen et al., 2005; Bestelmeyer & Mühl, 2022; Botvinick et al., 2004; Duggirala et al., 2022; Jardri et al., 2007). These regions may serve as a convergence zone for integrating identity-based and action-based information during speech monitoring.

Note that, although our paradigm was designed to examine sensory attenuation effects, we did not observe the expected suppression of the N1 or P2 for self-related voice input. Instead, both components showed enhanced amplitudes for self-voice and self-chosen words. This pattern diverges from traditional self-monitoring studies reporting, for example, N1 suppression for self-generated voice (Behroozmand et al., 2009; Heinks-Maldonado et al., 2007). This suggests that under conditions where predictive mechanisms are weakened (e.g., due to a delay between action and auditory feedback), memory-based identity recognition and attentional prioritization may become more dominant mechanisms for monitoring the self.

4.4. Limitations and Future Directions

The use of prerecorded voice feedback omitted bone-conducted components of natural speech (Reinfeldt et al., 2010; Orepic et al., 2023), which may contribute to the commonly reported experience of estrangement when hearing one's recorded voice (Daryadar & Raghibi, 2015). However, this effect may be waning in digitally literate populations (Thomas et al., 2024).

Although eLORETA provides robust estimates of intracranial sources (Mulert et al., 2004; Olbrich et al., 2009), it remains limited by the inverse problem: different source configurations can have identical scalp distributions (Grech et al., 2008). Multimodal approaches combining EEG with fMRI (e.g., EEG-fMRI fusion) could overcome this limitation.

Future work should systematically examine other agency dimensions, such as temporal contiguity, and distinguish between feeling and judgement of agency (Synofzik et al., 2008). Moreover, including passive and real-time feedback conditions, as common in classic self-stimulation paradigms, would help isolate sensorimotor from memory-based components of self-monitoring.

Clinically, this paradigm could inform the development of EEG-based diagnostic tools to stratify individuals according to specific disruptions in self-related processing. Such tools could help disentangle overlapping symptom profiles and support longitudinal monitoring of self-processing in psychiatric and neurodevelopmental populations.

5. Conclusion

This study provides novel evidence for a functional and temporal dissociation between voice ownership and agency during speech monitoring. N1 amplitude and frontotemporal-insular sources indexed voice ownership, reflecting rapid recognition of self-related identity cues. In contrast, P2 amplitude and parietal-sensory sources reflected agency, indexing the evaluation of action outcomes based on internal goals. The interaction between these dimensions in the P2 time window suggests that self-related processes, while dissociable, also integrate dynamically when monitoring speech feedback.

Together, these findings support a hybrid model of self-voice processing in which the forward model integrates both sensorimotor predictions and memory-based identity templates. This framework advances current theories of self-monitoring and has potential translational value. By delineating distinct neural markers of agency and ownership, this work lays the foundation for EEG-based tools that could aid in the identification and tracking of self-processing disruptions in clinical populations.

Acknowledgements

The authors would like to thank all the participants that completed the study and Diego Pinal for providing detailed tutorials on the use of eLORETA for source localization analysis, as well as for insightful discussions regarding related data analyses.

Funding sources

This work was supported by BIAL Foundation (BIAL 146/2020) and FCT - Fundação para a Ciência e Tecnologia (https://doi.org/10.54499/2022.13891.BD).

Credit author statement

APP, SAK, and MM conceptualized the study and developed the methodology. MM and MA collected and analyzed the data. MM undertook the statistical analyses. MM and APP drafted the original manuscript. APP and SAK contributed to the review and editing of the manuscript. All authors have read and approved the final manuscript.

Declaration of competing interest

The authors declare that they have no competing interests.

Data availability

The conditions of our ethics approval do not permit sharing of the data supporting the conclusions in this study with any individual outside the author team under any circumstances. These restrictions also apply to the study materials since they include self-voice recordings. However, the scripts used for the experimental tasks and statistical analyses, including the parameter files for source analysis tests, are available at https://osf.io/n9jxa/.

References

Allen, P. P., Amaro, E., Fu, C. H., Williams, S. C., Brammer, M., Johns, L. C., & McGuire, P. K. (2005). Neural correlates of the misattribution of self-generated speech. *Human brain mapping*, *26*(1), 44-53. https://doi.org/10.1002/hbm.20120

Allen, P., Freeman, D., Johns, L., & McGuire, P. (2006). Misattribution of self-generated speech in relation to hallucinatory proneness and delusional ideation in healthy volunteers. *Schizophrenia research*, *84*(2-3), 281-288. https://doi.org/10.1016/j.schres.2006.01.021

Andics, A., McQueen, J. M., & Petersson, K. M. (2013). Mean-based neural coding of voices. *Neuroimage*, *79*, 351-360. https://doi.org/10.1016/j.neuroimage.2013.05.002

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, *59*(4), 390-412. https://doi.org/10.1016/j.jml.2007.12.005

Bagiella, E., Sloan, R. P., & Heitjan, D. F. (2000). Mixed-effects models in psychophysiology. *Psychophysiology*, *37*(1), 13-20. https://doi.org/10.1111/1469-8986.3710013

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of statistical software*, *67*, 1-48. https://doi.org/10.18637/jss.v067.i01

Behroozmand, R., Karvelis, L., Liu, H., & Larson, C. R. (2009). Vocalization-induced enhancement of the auditory cortex responsiveness during voice F0 feedback perturbation. *Clinical neurophysiology*, *120*(7), 1303-1312. https://doi.org/10.1016/j.clinph.2009.04.022

Behroozmand, R., & Larson, C. R. (2011). Error-dependent modulation of speech-induced auditory suppression for pitch-shifted voice feedback. *BMC neuroscience*, *12*, 1-10. https://doi.org/10.1186/1471-2202-12-54

Beño-Ruiz-de-la-Sierra, R. M., Arjona-Valladares, A., Fondevila Estevez, S., Fernández-Linsenbarth, I., Díez, Á., & Molina, V. (2023). Corollary discharge function in healthy controls: Evidence about self-speech and external speech processing. *European Journal of Neuroscience*, *58*(7), 3705-3713. https://doi.org/10.1111/ejn.16125

Beño-Ruiz-de-la-Sierra, R. M., Arjona-Valladares, A., Hernández-García, M., Fernández-Linsenbarth, I., Díez, Á., Fondevila Estevez, S., Castaño, C., Muñoz, F., Sanz-Fuentenebro, J., Roig-Herrero, A., & Molina, V. (2024). Corollary discharge dysfunction as a possible substrate of anomalous self-experiences in schizophrenia. *Schizophrenia Bulletin*, *50*(5), 1137-1146. https://doi.org/10.1093/schbul/sbad157

Bestelmeyer, P. E., & Mühl, C. (2022). Neural dissociation of the acoustic and cognitive representation of voice identity. *NeuroImage*, *263*, 119647. https://doi.org/10.1016/j.neuroimage.2022.119647

Boisgontier, M. P., & Cheval, B. (2016). The anova to mixed model transition. *Neuroscience & Biobehavioral Reviews*, *68*, 1004-1005. https://doi.org/10.1016/j.neubiorev.2016.05.034

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in cognitive sciences*, *8*(12), 539-546. https://doi.org/10.1016/j.tics.2004.10.003

Bühler, T., Kindler, J., Schneider, R. C., Strik, W., Dierks, T., Hubl, D., & Koenig, T. (2016). Disturbances of agency and ownership in schizophrenia: An auditory verbal event related potentials study. *Brain topography*, *29*, 716-727. https://doi.org/10.1007/s10548-016-0495-1

Chambon, V., Moore, J. W., & Haggard, P. (2015). TMS stimulation over the inferior parietal cortex disrupts prospective sense of agency. *Brain Structure and Function*, *220*(6), 3627-3639. https://doi.org/10.1007/s00429-014-0878-6

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: Age, sleep and modality. *Clinical Neurophysiology, 115*(4), 732e744. https:// doi.org/10.1016/j.clinph.2003.11.021

Daryadar, M., & Raghibi, M. (2015). The effect of listening to recordings of one's voice on attentional bias and auditory verbal learning. *International Journal of Psychological Studies*, *7*(2), 155-163. https://doi.org/10.5539/ijps.v7n2p155

Desantis, A., Weiss, C., Schütz-Bosbach, S., & Waszak, F. (2012). Believing and perceiving: authorship belief modulates sensory attenuation. *PLoS One*, *7*(5), e37959. https://doi.org/10.1371/journal.pone.0037959

Di Plinio, S., Perrucci, M. G., Aleman, A., & Ebisch, S. J. H. (2020). I am Me: Brain systems integrate and segregate to establish a multidimensional sense of self. *NeuroImage*, *205*, 116284. https://doi.org/10.1016/j.neuroimage.2019.116284

Duggirala, S. X., Belyk, M., Schwartze, M., Kanske, P., & Kotz, S. A. (2022). Emotional salience but not valence impacts anterior cingulate cortex conflict processing. *Cognitive, Affective, & Behavioral Neuroscience*, *22*(6), 1250-1263. https://doi.org/10.3758/s13415-022-01025-9

Duzcu, H., Ozkurt, T. E., Mapelli, I., & Hohenberger, A. (2019). N1-P2: Neural markers of temporal expectation and response discrimination in interval timing. *Acta Neurobiologiae Experimentalis, 79*(2), 193e204. https://doi.org/10.21307/ane2019-017

Egan, S., Ghio, M., & Bellebaum, C. (2023). Auditory N1 and P2 attenuation in action observation: An event-related potential study considering effects of temporal predictability and individualism. *Biological Psychology*, *180*, 108575. https://doi.org/10.1016/j.biopsycho.2023.108575

Esslen, M., Metzler, S., Pascual-Marqui, R., & Jancke, L. (2008). Pre-reflective and reflective self-reference: A spatiotemporal EEG analysis. *NeuroImage*, *42*(1), 437–449. https://doi.org/10.1016/j.neuroimage.2008.01.060

Fu, C. H., Vythelingum, G. N., Brammer, M. J., Williams, S. C., Amaro Jr, E., Andrew, C. M., Yáguez, L., van Haren, N. E. M., Matsumoto, K., & McGuire, P. K. (2006). An fMRI study of verbal self-monitoring: neural correlates of auditory verbal feedback. *Cerebral cortex*, *16*(7), 969-977. https://doi.org/10.1093/cercor/bhj039

Gentsch, A., & Schütz-Bosbach, S. (2011). I did it: unconscious expectation of sensory consequences modulates the experience of self-agency and its functional signature. *Journal of cognitive neuroscience*, *23*(12), 3817-3828. https://doi.org/10.1162/jocn_a_00012

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and clinical neurophysiology*, *55*(4), 468-484. https://doi.org/10.1016/0013-4694(83)90135-9

Grech, R., Cassar, T., Muscat, J., Camilleri, K. P., Fabri, S. G., Zervakis, M., Xanthopoulos, P., Sakkalis, V., & Vanrumste, B. (2008). Review on solving the inverse problem in EEG source analysis. *Journal of neuroengineering and rehabilitation*, *5*(1), 25. https://doi.org/10.1186/1743-0003-5-25

Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, *18*(4), 196-207. https://doi.org/10.1038/nrn.2017.14

Han, N., Jack, B. N., Hughes, G., Elijah, R. B., & Whitford, T. J. (2021). Sensory attenuation in the absence of movement: Differentiating motor action from sense of agency. *Cortex*, *141*, 436-448. https://doi.org/10.1016/j.cortex.2021.04.010

Heinks-Maldonado, T. H., Mathalon, D. H., Houde, J. F., Gray, M., Faustman, W. O., & Ford, J. M. (2007). Relationship of imprecise corollary discharge in schizophrenia to auditory hallucinations. *Archives of general psychiatry*, *64*(3), 286-296. https://doi.org/10.1001/archpsyc.64.3.286

Hubl, D., Schneider, R. C., Kottlow, M., Kindler, J., Strik, W., Dierks, T., & Koenig, T. (2014). Agency and ownership are independent components of 'sensing the self' in the auditory-verbal domain. *Brain topography*, *27*, 672-682. https://doi.org/10.1007/s10548-014-0351-0

Iannotti, G. R., Orepic, P., Brunet, D., Koenig, T., Alcoba-Banqueri, S., Garin, D. F., Schaller, K., Blanke, O., & Michel, C. M. (2022). EEG spatiotemporal patterns underlying self-other voice discrimination. *Cerebral cortex*, *32*(9), 1978-1992. https://doi.org/10.1093/cercor/bhab329

Jardri, R., Pins, D., Bubrovszky, M., Despretz, P., Pruvo, J. P., Steinling, M., & Thomas, P. (2007). Self awareness and speech processing: an fMRI study. *Neuroimage*, *35*(4), 1645-1653. https://doi.org/10.1016/j.neuroimage.2007.02.002

Johns, L. C., Allen, P., Valli, I., Winton-Brown, T., Broome, M., Woolley, J., Tabraham, P., Day, F., Howes, O., Wykes, T., & McGuire, P. (2010). Impaired verbal self-monitoring in individuals at high risk of psychosis. *Psychological Medicine*, *40*(9), 1433-1442. https://doi.org/10.1017/S0033291709991991

Johnson, J. F., Belyk, M., Schwartze, M., Pinheiro, A. P., & Kotz, S. A. (2021). Expectancy changes the self-monitoring of voice identity. *European Journal of Neuroscience*, *53*(8), 2681-2695. https://doi.org/10.1111/ejn.15162

Kalckert, A., & Henrik Ehrsson, H. (2012). Moving a rubber hand that feels like your own: A dissociation of ownership and agency. *Frontiers in Human Neuroscience*, *6*, 40. https://doi.org/10.3389/fnhum.2012.00040

Kawahara, H., & Morise, M. (2011). Technical foundations of TANDEM-STRAIGHT, a speech analysis, modification and synthesis framework. *Sadhana*, *36*, 713-727. https://doi.org/10.1007/s12046-011-0043-3

Knolle, F., Schröger, E., Baess, P., & Kotz, S. A. (2012). The cerebellum generates motor-to-auditory predictions: ERP lesion evidence. *Journal of cognitive neuroscience*, *24*(3), 698-706. https://doi.org/10.1162/jocn_a_00167

Knolle, F., Schröger, E., & Kotz, S. A. (2013). Prediction errors in self-and externally-generated deviants. *Biological Psychology*, *92*(2), 410-416. https://doi.org/10.1016/j.biopsycho.2012.11.017

Knolle, F., Schwartze, M., Schröger, E., & Kotz, S. A. (2019). Auditory predictions and prediction errors in response to self-initiated vowels. *Frontiers in neuroscience*, *13*, 1146. https://doi.org/10.3389/fnins.2019.01146

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *Journal of statistical software*, *82*, 1-26. https://doi.org/10.18637/jss.v082.i13

Lavan, N. (2023). How do we describe other people from voices and faces?. *Cognition*, *230*, 105253. https://doi.org/10.1016/j.cognition.2022.105253

Latinus, M., & Belin, P. (2011). Human voice perception. *Current Biology*, *21*(4), 143-145. https://doi.org/10.1016/j.cub.2010.12.033

Liu, M., He, X., Rotsthein, P., & Sui, J. (2016). Dynamically orienting your own face facilitates the automatic attraction of attention. *Cognitive neuroscience*, *7*(1-4), 37-44. https://doi.org/10.1080/17588928.2015.1044428

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology*, *54*(1), 146-157. https://doi.org/10.1111/psyp.12639

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior research methods*, *49*, 1494-1502. https://doi.org/10.3758/s13428-016-0809-y

Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural networks*, *9*(8), 1265-1279. https://doi.org/10.1016/S0893-6080(96)00035-4

Mulert, C., Jäger, L., Schmitt, R., Bussfeld, P., Pogarell, O., Möller, H. J., Juckel, G. & Hegerl, U. (2004). Integration of fMRI and simultaneous EEG: towards a comprehensive understanding of localization and time-course of brain activity in target detection. *Neuroimage*, *22*(1), 83-94. https://doi.org/10.1016/j.neuroimage.2003.10.051

Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human brain mapping*, *15*(1), 1-25. https://doi.org/10.1002/hbm.1058

Niziolek, C. A., Nagarajan, S. S., & Houde, J. F. (2013). What does motor efference copy represent? Evidence from speech production. *Journal of Neuroscience*, *33*(41), 16110-16116. https://doi.org/10.1523/JNEUROSCI.2137-13.2013

Ohata, R., Asai, T., Imaizumi, S., & Imamizu, H. (2022). I hear my voice; therefore I spoke: The sense of agency over speech is enhanced by hearing one's own voice. *Psychological science*, *33*(8), 1226-1239. https://doi.org/10.1177/09567976211068880

Olbrich, S., Mulert, C., Karch, S., Trenner, M., Leicht, G., Pogarell, O., & Hegerl, U. (2009). EEG-vigilance and BOLD effect during simultaneous EEG/fMRI measurement. *neuroimage*, *45*(2), 319-332. https://doi.org/10.1016/j.neuroimage.2008.11.014

Orepic, P., Kannape, O. A., Faivre, N., & Blanke, O. (2023). Bone conduction facilitates self-other voice discrimination. *Royal Society Open Science*, *10*(2), 221561. https://doi.org/10.1098/rsos.221561

Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. *Methods find exp clin pharmacol*, *24*(Suppl D), 5-12.

Pascual-Marqui, R. D. (2007). Discrete, 3D distributed, linear imaging methods of electric neuronal activity. Part 1: exact, zero error localization. *arXiv preprint arXiv:0710.3341*.

Picard, F., & Friston, K. (2014). Predictions, perception, and a sense of self. *Neurology*, *83*(12), 1112-1118. https://doi.org/10.1212/WNL.0000000000000798

Pinal, D., Zurrón, M., & Díaz, F. (2014). Effects of load and maintenance duration on the time course of information encoding and retrieval in working memory: from perceptual analysis to post-categorization processes. *Frontiers in human neuroscience*, 165(*8)*,1-15. https://doi.org/10.3389/fnhum.2014.00165

Pinheiro, A. P., Rezaii, N., Nestor, P. G., Rauber, A., Spencer, K. M., & Niznikiewicz, M. (2016). Did you or I say pretty, rude or brief? An ERP study of the effects of speaker's identity on emotional word processing. *Brain and Language*, *153–154*, 38–49. https://doi.org/10.1016/j.bandl.2015.12.003

Pinheiro, A. P., Farinha-Fernandes, A., Roberto, M. S., & Kotz, S. A. (2019a). Self-voice perception and its relationship with hallucination predisposition. *Cognitive neuropsychiatry*, *24*(4), 237-255. https://doi.org/10.1080/13546805.2019.1621159

Pinheiro, A. P., Schwartze, M., Gutierrez, F., & Kotz, S. A. (2019b). When temporal prediction errs: ERP responses to delayed action-feedback onset. *Neuropsychologia*, *134*, 107200. https://doi.org/10.1016/j.neuropsychologia.2019.107200

Pinheiro, A. P., Schwartze, M., Amorim, M., Coentre, R., Levy, P., & Kotz, S. A. (2020). Changes in motor preparation affect the sensory consequences of voice production in voice hearers. *Neuropsychologia*, *146*, 107531. https://doi.org/10.1016/j.neuropsychologia.2020.107531

Pinheiro, A. P., Sarzedas, J., Roberto, M. S., & Kotz, S. A. (2023). Attention and emotion shape self-voice prioritization in speech processing. *Cortex*, *158*, 83-95. https://doi.org/10.1016/j.cortex.2022.10.006

Poonian, S. K., McFadyen, J., Ogden, J., & Cunnington, R. (2015). Implicit agency in observed actions: evidence for N1 suppression of tones caused by self-made and observed actions. *Journal of cognitive neuroscience*, *27*(4), 752-764. https://doi.org/10.1162/jocn_a_00745

Reinfeldt, S., Östli, P., Håkansson, B., & Stenfelt, S. (2010). Hearing one's own voice during phoneme vocalization - Transmission by air and bone conduction. *The Journal of the Acoustical Society of America*, *128*(2), 751-762. https://doi.org/10.1121/1.3458855

Saha, S., Nesterets, Y. I., Tahtali, M., & Gureyev, T. E. (2015). Evaluation of spatial resolution and noise sensitivity of sLORETA method for EEG source localization using low-density headsets. *Biomedical physics & engineering express*, *1*(4), 045206. https://doi.org/10.1088/2057-1976/1/4/045206

Schröger, E., Kotz, S. A., & SanMiguel, I. (2015). Bridging prediction and attention in current research on perception and action. *Brain Research*, *1626*, 1-13. https://doi.org/10.1016/j.brainres.2015.08.037

Seghezzi, S., Zirone, E., Paulesu, E., & Zapparoli, L. (2019). The brain in (willed) action: a meta-analytical comparison of imaging studies on motor intentionality and sense of agency. *Frontiers in psychology*, *10*, 804. https://doi.org/10.3389/fpsyg.2019.00804

Seidel, A., Ghio, M., Studer, B., & Bellebaum, C. (2021). Illusion of control affects ERP amplitude reductions for auditory outcomes of self-generated actions. *Psychophysiology*, *58*(5), e13792. https://doi.org/10.1111/psyp.13792

Soares, A. P., Comesaña, M., Pinheiro, A. P., Simões, A., & Frade, C. S. (2012). The adaptation of the Affective Norms for English words (ANEW) for European Portuguese. *Behavior research methods*, *44*, 256-269. https://doi.org/10.3758/s13428-011-0131-7

Sui, J., He, X., Golubickis, M., Svensson, S. L., & Macrae, C. N. (2023). Electrophysiological correlates of self-prioritization. *Consciousness and Cognition*, *108*, 103475. https://doi.org/10.1016/j.concog.2023.103475

Synofzik, M., Vosgerau, G., & Newen, A. (2008). I move, therefore I am: A new theoretical framework to investigate agency and ownership. *Consciousness and Cognition*, *17*, 411–424. https://doi.org/10.1016/j.concog.2008.03.008

Takahashi, H., Rissling, A. J., Pascual-Marqui, R., Kirihara, K., Pela, M., Sprock, J., Braff, D. L., & Light, G. A. (2013). Neural substrates of normal and impaired preattentive sensory discrimination in large cohorts of nonpsychiatric subjects and schizophrenia patients as indexed by MMN and P3a change detection responses. *Neuroimage*, *66*, 594-603. https://doi.org/10.1016/j.neuroimage.2012.09.074

Thomas, J., Jogia, J., Barbato, M., & Bentall, R. (2024). Me, not-me: Voice note use predicts self-voice recognition and liking. *Computers in Human Behavior Reports*, *15*, 100446.https://doi.org/10.1016/j.chbr.2024.100446

Timm, J., Schönwiesner, M., Schröger, E., & SanMiguel, I. (2016). Sensory suppression of brain responses to self-generated sounds is observed with and without the perception of agency. *Cortex*, *80*, 5-20. https://doi.org/10.1016/j.cortex.2016.03.018

Uhlmann, L., Pazen, M., van Kemenade, B. M., Kircher, T., & Straube, B. (2021). Neural correlates of self-other distinction in patients with schizophrenia spectrum disorders: The roles of agency and hand identity. *Schizophrenia bulletin*, *47*(5), 1399-1408. https://doi.org/10.1093/schbul/sbaa186

Wagner, M., Shafer, V. L., Haxhari, E., Kiprovski, K., Behrmann, K., & Griffiths, T. (2017). Stability of the cortical sensory waveforms, the P1-N1-P2 complex and T-complex, of auditory evoked potentials. *Journal of Speech, Language, and Hearing Research, 60*(7), 2105-2115. https://doi.org/10.1044/2017_JSLHR-H-16-0056

Weiss, C., Herwig, A., & Schütz-Bosbach, S. (2011). The self in action effects: selective attenuation of self-generated sounds. *Cognition*, *121*(2), 207-218. https://doi.org/10.1016/j.cognition.2011.06.011

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*, 644059. https://doi.org/10.3389/fnsys.2021.644059

Wolpert, D. M. (1997). Computational approaches to motor control. *Trends in cognitive sciences*, *1*(6), 209-216. https://doi.org/10.1016/S1364-6613(97)01070-X

Zhao, W., Luo, L., Li, Q., & Kendrick, K. M. (2013). What can psychiatric disorders tell us about neural processing of the self?. *Frontiers in Human Neuroscience*, *485*(7), 1-15. https://doi.org/10.3389/fnhum.2013.00485

Figures

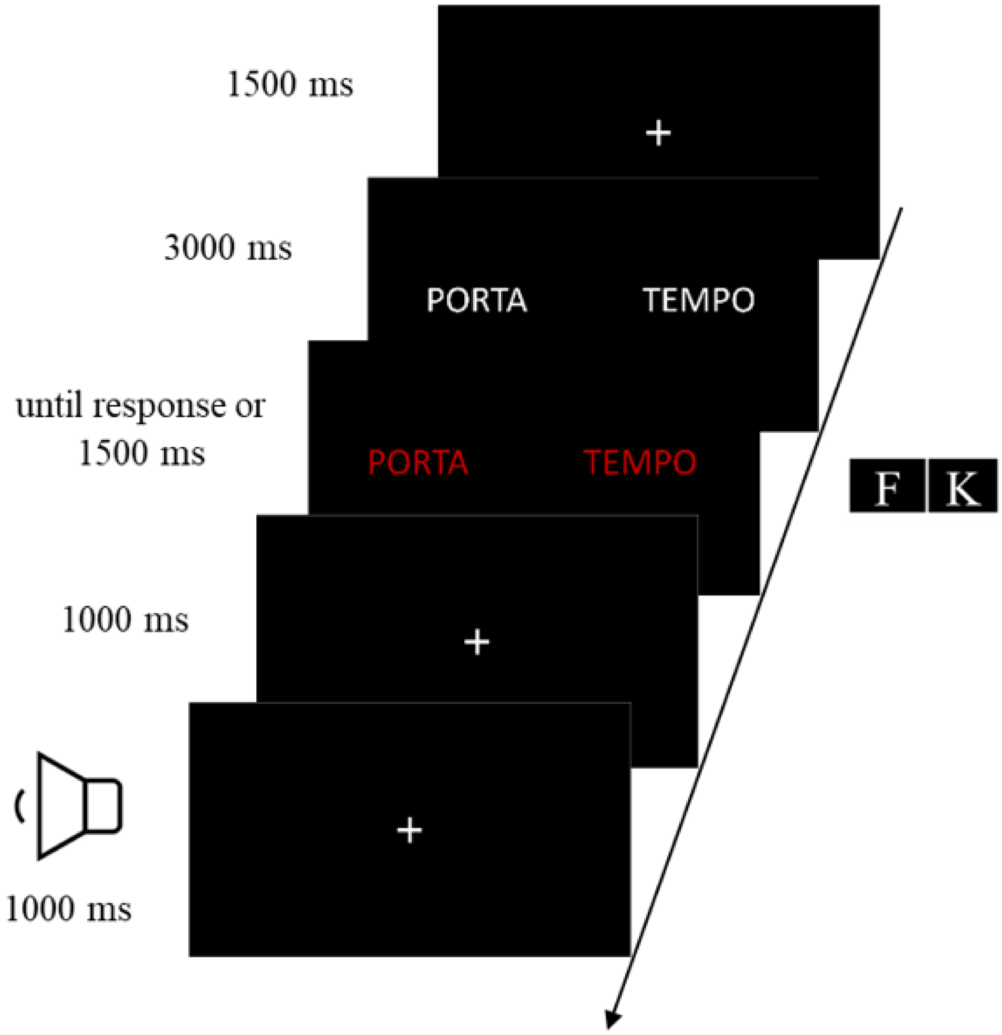

Figure 1. Trial structure of the paradigm adapted from Di Plinio et al. (2020). ‘Porta’ (‘door’) and ‘Tempo’ (‘time’) are an example of a word pair used in the study.

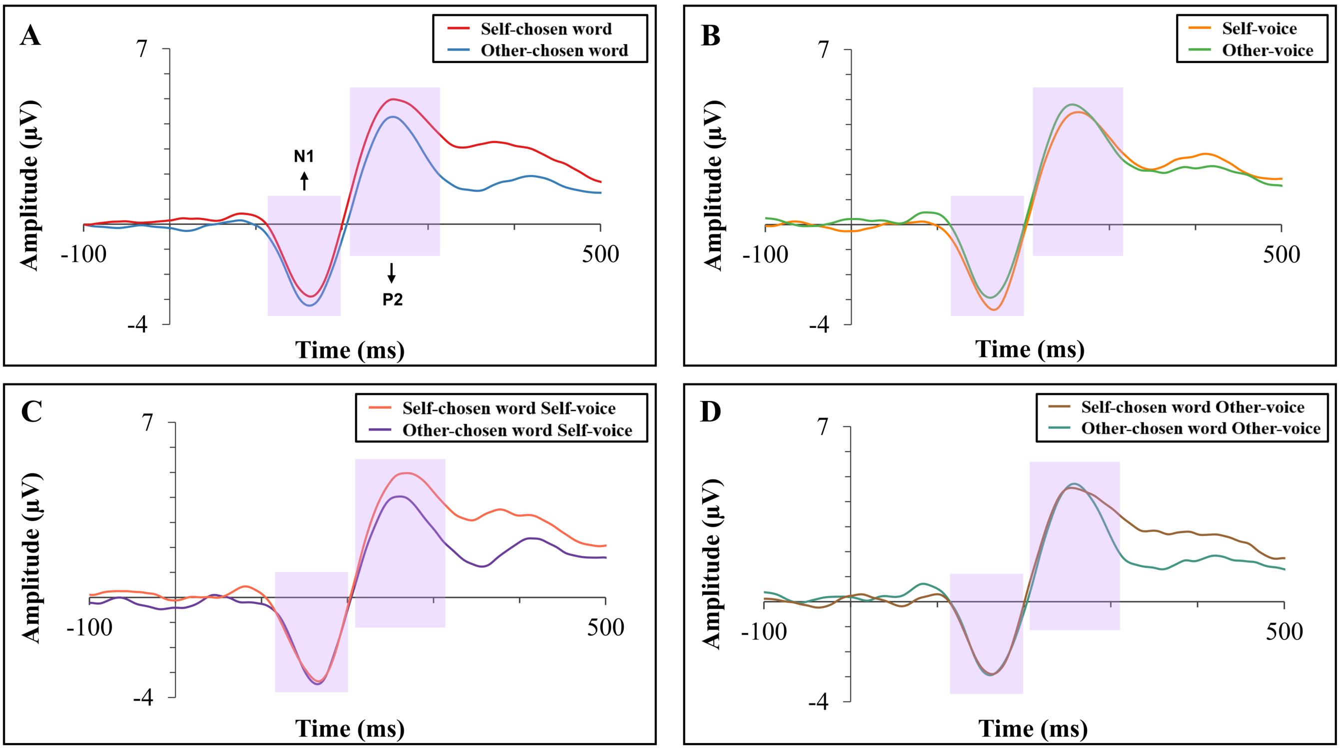

Figure 2. Grand average (GA) waveforms illustrating the N1 and P2 responses, collapsed across frontal (F3, Fz, F4), frontocentral (FC3, FCz, FC4), and central (C3, Cz, C4) electrodes. **A.** Self-chosen and other-chosen word conditions. **B.** Self-voice and other-voice conditions. **C** and **D.** Illustrate the statistically significant interaction between agency and ownership found in the P2 time-window.

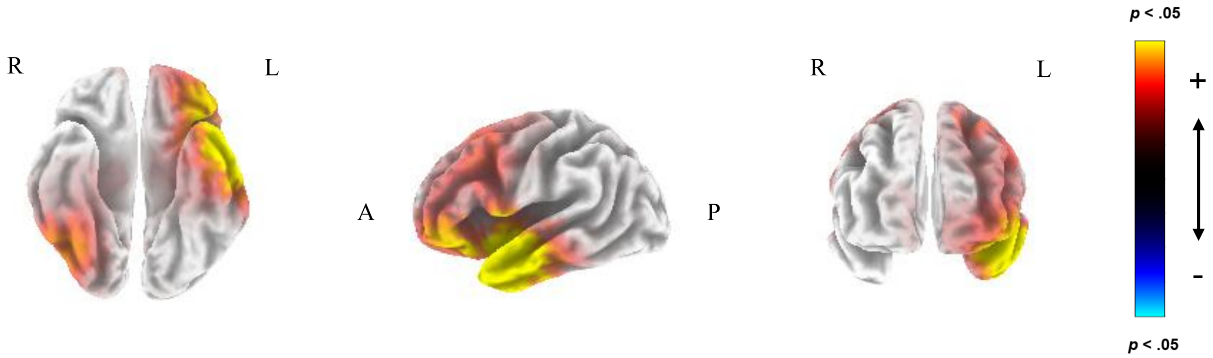

Figure 3. Three dimensional eLORETA statistical maps for the self-voice and other-voice contrast: bottom, left, frontal view. Color indicates voxel-by-voxel LnF values (p = .05 for threshold LnF value of .540). Significant differences in activation (p < .05) are color coded as bright yellow. L, left; R, right; A, anterior; P, posterior.

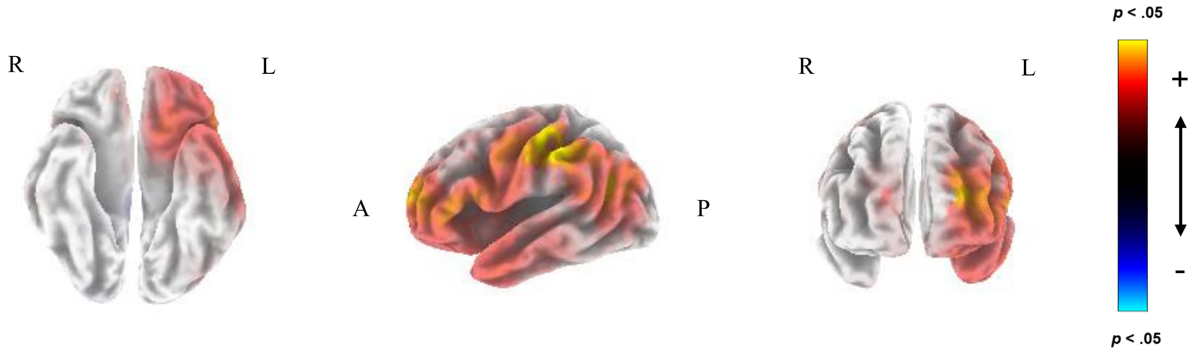

Figure 4. Three dimensional eLORETA statistical maps for the self-chosen and other-chosen word contrast: bottom, left, frontal view. Color indicates voxel-by-voxel LnF values (p = .05 for threshold LnF value of .364). Significant differences in activation ($p < .05$) are color coded as bright yellow. L, left; R, right; A, anterior; P, posterior.

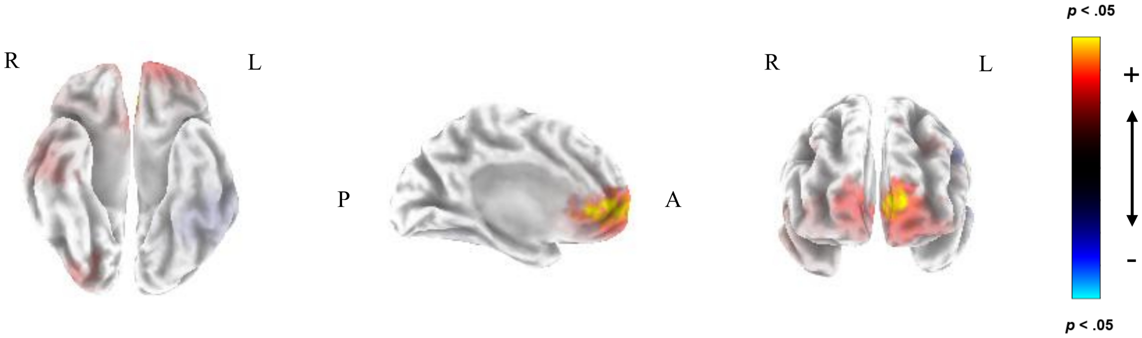

Figure 5. Three dimensional eLORETA statistical maps for the contrast between (SS – OS) and (SO – OO): bottom, left, frontal view. Color indicates voxel-by-voxel t-statistics (p = .05 for t-value of 3.486). Significant differences in activation (p < .05) are color coded as bright yellow. L, left; R, right; A, anterior; P, posterior.

Tables

Table 1

eLORETA results for the contrast Self-voice vs. Other-voice overlapping with the N1 time window

Lobe	Structure	BA	MNI coordinates			N of voxels	Peak voxel value
			x	y	z		
Temporal	Middle temporal gyrus	21	-50	0	15	40	.616
	Superior temporal gyrus	21/ 22/ 38	-50	5	-15	57	.615
	Inferior temporal gyrus	20/ 21	-60	-5	-20	11	.609
	Sub-Gyral	13/ 21	-40	0	-10	3	.588
Sub-lobar	Insula	13	-40	0	-5	5	.572
Frontal	Inferior frontal gyrus	47	-40	15	-15	3	.551

Note: BA corresponds to Brodmann Areas. MNI coordinates indicate the location of peak voxel activation.

Table 2

eLORETA results for the contrast Self-chosen vs. Other-chosen word overlapping with the P2 time window

Lobe	Structure	BA	MNI coordinates			N of voxels	Peak voxel value
			x	y	z		
Parietal	Inferior parietal lobule	40	-50	-30	45	5	.393
	Postcentral gyrus	1/ 2/ 3	-55	-25	50	17	.388
	Angular gyrus	39	-30	-60	25	3	.377

Table 3

eLORETA results for the contrast (SS – OS) vs. (SO – OO) overlapping with the P2 time window

Lobe	Structure	BA	MNI coordinates			N of voxels	Peak voxel value
			x	y	z		
Limbic	Anterior Cingulate gyrus	32	-10	45	5	2	3.728
Frontal	Medial Frontal Gyrus	10	-10	55	5	3	3.644