Do Corollary Discharges Contain Information About the Volume of Inner Speech? An ERP Study

Kevin Berryman[1,2], Thomas J. Whitford[3], Mike E. Le Pelley[3], and Bradley N. Jack[1]

Abstract

■ When we move our articulator organs to produce overt speech, the brain generates a corollary discharge that acts to suppress the neural and perceptual responses to our speech sounds. Recent research shows that inner speech—the silent production of words in one's mind—is accompanied by a corollary discharge that contains information about the timing and content of inner speech. The aim of the present study was to determine whether this corollary discharge contains information about the volume of inner speech. To investigate this, participants watched an animation that provided them with precise knowledge about when they should produce a "loud" or "quiet" inner sound. At the same time, they heard an audible sound that was either similar or dissimilar to the volume of the inner sound. We found that producing the inner sound attenuated the N1—an ERP signature of auditory processing—and enhanced the slow negative wave—an ERP signature of anticipation and motor preparation—compared with listening, regardless of whether the volume of the inner sound was similar or dissimilar to the audible sound. We also found that the volume of the inner sound did not differentially modulate the N1 or slow negative wave. We speculate that this might be because one of the functions of corollary discharge is to protect our auditory receptors from desensitization caused by audible sounds, and this might be redundant in the context of inner speech because inner speech does not produce an audible sound. We conclude that there might be a functional difference between the neural processes that underlie the production of inner and overt speech. ■

INTRODUCTION

When we produce overt speech, the sound of our voice elicits smaller neural and perceptual responses than when the same sounds are replayed to us. This phenomenon is known as sensory suppression (Horváth, 2015; Schröger, Marzecová, & SanMiguel, 2015; Hughes, Desantis, & Waszak, 2013; Bendixen, SanMiguel, & Schröger, 2012) and is thought to result from the action of an internal forward model (Wolpert & Miall, 1996). According to this framework, when we move our articulator organs to speak, an efference copy of the motor command (Von Holst & Mittelstaedt, 1950) is used to compute a neural prediction—a corollary discharge (Sperry, 1950)—of our speech sounds. If the corollary discharge matches our actual speech sounds, neural and perceptual responses are suppressed or "explained away"; if the corollary discharge does not match our speech sounds, the difference between them is sent for further processing (Straka, Simmers, & Chagnaud, 2018; Crapse & Sommer, 2008; Poulet & Hedwig, 2007; Schütz-Bosbach & Prinz, 2007). There is a large body of research showing that corollary discharges associated with overt speech contain information about the content, timing, and volume of overt speech (Behroozmand, Sangtian, Korzyukov, & Larson, 2016; Sitek et al., 2013; Chen, Chen, Liu, Huang, & Liu, 2012; Behroozmand & Larson, 2011; Behroozmand, Liu, & Larson, 2011; Liu, Meshman, Behroozmand, & Larson, 2011; Behroozmand, Karvelis, Liu, & Larson, 2009; Eliades & Wang, 2008; Bauer, Mittal, Larson, & Hain, 2006; Heinks-Maldonado & Houde, 2005; Heinks-Maldonado, Mathalon, Gray, & Ford, 2005; Houde, Nagarajan, Sekihara, & Merzenich, 2002). Recent research has shown that corollary discharges are also associated with inner speech—the silent production of words in one's mind (Alderson-Day & Fernyhough, 2015; Perrone-Bertolotti, Rapin, Lachaux, Baciu, & Lœvenbruck, 2014; Zivin, 1979)—and that they contain information about the content and timing of inner speech (Whitford et al., 2017, 2025; Chung et al., 2023; Jack et al., 2019; Tian, Ding, Teng, Bai, & Poeppel, 2018; Tian, Zarate, & Poeppel, 2016; Tian & Poeppel, 2010, 2012, 2013, 2015; Ylinen et al., 2015; Scott, 2013; Ford & Mathalon, 2004). The aim of the present study was to determine whether corollary discharges contain information about the volume of inner speech.

As mentioned above, recent research has shown that corollary discharges are associated with inner speech. For instance, Whitford and colleagues (2017) introduced a procedure in which participants viewed a ticker-tape-style cue, which provided them with precise knowledge about when they would hear an audible sound. In the *listen condition* of their experiment, participants were instructed to passively listen to the audible sound; in the *inner speech condition*, participants were instructed to

[1]Australian National University, [2]Monash University, [3]University of New South Wales

https://doi.org/10.1162/jocn_a_02318

produce an inner sound at the precise moment they heard the audible sound. On a random half of the trials in the inner speech condition, the inner and audible sounds matched on content—this was called the *match condition*; on the other half of the trials, the inner and audible sounds did not match on content—this was called the *mismatch condition*. Jack and colleagues (2019) used a similar procedure, except that they also presented the audible sound 300 msec before, concurrently with, or 300 msec after participants produced the inner sound. Both studies found that producing the inner sound attenuated the N1—an ERP signature of auditory processing (Woods, 1995; Näätänen & Picton, 1987)—compared with listening, but only when the inner and audible sounds occurred concurrently and matched on content. If the audible sound was presented before or after the production of the inner sound, or if the inner sound did not match the content of the audible sound, there was no attenuation of the N1. Chung and colleagues (2023) re-analyzed the Whitford and colleagues (2017) data and found that producing the inner sound enhanced the slow negative wave (SNW)—an ERP signature of anticipation and motor preparation (Schurger, Hu, Pak, & Roskies, 2021; Brunia, van Boxtel, & Böcker, 2012; van Boxtel & Böcker, 2004)—compared with listening. Several studies interpret the SNW as being evidence of corollary discharge (Ody, Kircher, Straube, & He, 2023; Pinheiro, Schwartze, Gutiérrez-Domínguez, & Kotz, 2020; Reznik, Simon, & Mukamel, 2018; Vercillo, O'Neil, & Jiang, 2018; Wen et al., 2018; Jo, Wittmann, Hinterberger, & Schmidt, 2014). These results suggest that inner speech, similar to overt speech, is accompanied by a content-specific and temporally precise corollary discharge, in that it contains information about the content and timing of inner speech.

However, when we move our articulator organs to produce overt speech, the accompanying corollary discharge is not just content-specific and temporally precise (Behroozmand et al., 2009, 2011, 2016; Sitek et al., 2013; Chen et al., 2012; Behroozmand & Larson, 2011; Liu et al., 2011; Eliades & Wang, 2008; Heinks-Maldonado et al., 2005; Houde et al., 2002) but also volume-dependent, in that it contains information about the volume of overt speech (Bauer et al., 2006; Heinks-Maldonado & Houde, 2005). The aim of the present study was twofold. First, we sought to determine whether inner speech, similar to overt speech, is accompanied by a volume-dependent corollary discharge. On one hand, this possibility seems likely because the neural processes underlying inner speech are thought to be the same as those underlying overt speech (Jones & Fernyhough, 2007; Frith, 1987; Feinberg, 1978). On the other hand, one of the proposed functions of corollary discharge is to filter auditory input to protect our auditory receptors from desensitization caused by audible sounds (Crapse & Sommer, 2008), and this function might be redundant in the context of inner speech because inner speech does not produce an audible sound. To investigate this, we used a similar procedure to Whitford and colleagues (2017), except that our participants were instructed to produce either a loud or quiet inner sound that was either similar or dissimilar to the volume of the audible sound. Here, we use the terms "similar" and "dissimilar," rather than "match" and "mismatch," respectively, because it is difficult to know whether the volume of the inner sound precisely matches the volume of the audible sound or not. If corollary discharges contain information about the volume of inner speech, we hypothesize a larger N1-attenuation effect when the volume of the inner sound is similar to the audible sound compared with when it is not. Alternatively, if corollary discharges do not contain information about the volume of inner speech, we hypothesize similar N1-attenuation effects for the similar and dissimilar conditions. We also sought to replicate the results of Chung and colleagues (2023) by observing the SNW.

The second aim of the present study was to determine whether the volume of inner speech influences the N1-attenuation effect or the magnitude of the SNW. A similar question was addressed by Tian and colleagues (2018), who wondered whether the volume of inner speech influences subsequent loudness perception and auditory processing. To test this, they used a repetition procedure in which their participants produced either a loud or quiet inner sound multiple times, which was followed by an audible sound whose volume remained constant. In one of their behavioral experiments, they found that the perceived volume of the audible sound was rated as quieter after producing four loud inner sounds compared with producing four quiet inner sounds. In a different behavioral experiment, they found that the number of repetitions influenced loudness ratings: If the inner sound was produced three to six times, then the perceived volume of the audible sound was rated as quieter after producing loud inner sounds compared with producing quiet inner sounds, but if the inner sound was produced once or twice, then the perceived volume of the audible sound was rated similarly for loud and quiet inner sounds. Finally, in an ERP experiment, they found that producing four inner sounds attenuated the N1 compared with listening and that the N1-attenuation effect was larger for loud inner sounds compared with quiet inner sounds. These results suggest that the volume of inner speech influences subsequent loudness perception and auditory processing. By reconfiguring our repeated-measures ANOVA from listen versus similar versus dissimilar to listen versus loud inner sound versus quiet inner sound, we sought to conceptually replicate the N1 results of Tian and colleagues (2018), as well as extend upon them by investigating the SNW and determining whether it is also influenced by the volume of inner speech.

METHODS

Participants

Sixty-one students from the Australian National University (ANU) participated in our experiment for course credit. All participants gave written informed consent before the experiment and reported having normal hearing in both ears. Data from two participants were excluded from further analyses due to excessive artifacts in the EEG recording (>75% of epochs meeting the rejection criteria; see below). Mean age of the remaining participants, 33 of whom were female and 53 of whom were right-handed, was 21 (*SD* = 2) years. This sample size is similar to that used by Chung and colleagues (2023), Jack and colleagues (2019), and Whitford and colleagues (2017) and is consistent with recent attempts to estimate statistical power for ERP studies using the N1 (Hall et al., 2023). The experiment was approved by ANU's Science and Medical Delegated Ethics Review Committee and was conducted in accordance with the ethical standards laid down in the Declaration of Helsinki (World Medical Association, 2013).

Apparatus, Stimuli, and Procedure

Participants sat in a quiet, dimly lit room, approximately 60 cm in front of a computer monitor (Dell U2415) and wore headphones (Audio-Technica ATH-M20x). Stimulus presentation was controlled by specially written MATLAB (The MathWorks) scripts using the Psychophysics Toolbox (Kleiner, Brainard, & Pelli, 2007; Brainard, 1997; Pelli, 1997). Participants watched an animation, which was identical to the one used by Whitford and colleagues (2017), on every trial: It began with a green horizontal line in the center of the screen—the ticker-tape—a red vertical line in the center of the screen—the fixation line—and a green vertical line on the right-hand side of the screen—the target line (see Figure 1A). Participants were instructed to look at the fixation line for the duration of the trial. After a 1-sec delay, the target line began to move leftward across the screen at a speed of 6.5°/sec, such that after 4 sec the target line overlapped the fixation line and subsequently continued to move across the ticker-tape for an additional 1 sec (see Figure 1C–F). An auditory stimulus was presented at the precise moment the target and fixation lines overlapped (see Figure 1E). At the end of each trial, participants rated their subjective performance on that trial with a 5-point Likert scale, with scores ranging from 1, meaning *not at all successful*, to 5, meaning *completely successful*. We used these ratings to classify trials in which participants successfully performed the task or not.

The experiment consisted of six blocks of trials, with each block containing 60 trials. In two of the blocks, participants performed the listen condition: They were instructed to passively listen to a recording of the audible sound /ba/, which was produced by a male speaker and was about 200 msec long (see Figure 1G and J). In two different blocks, participants performed the loud inner sound condition: They listened to a recording of the audible sound /ba/, and they were instructed to loudly produce the sound /ba/ in their minds at the precise moment the fixation and target lines overlapped (see Figure 1H and L). In the remaining two blocks, participants performed the quiet inner sound condition: They listened to a recording of the audible sound /ba/, and they were instructed to quietly produce the sound /ba/ in their minds at the precise moment the fixation and target lines overlapped (see Figure 1I and K). The order of the blocks was counterbalanced.

On half of the trials in each block, the audible sound was 80 dB SPL; on the other half of trials, the audible sound was 60 dB. This meant that, on half of the trials in the inner speech blocks, the inner and audible sounds matched on volume; that is, participants loudly or quietly produced the sound /ba/ in their minds and listened to a loud or quiet recording of the sound /ba/, respectively—the similar condition (see Figure 1H and K). On the other half of trials, the inner and audible sounds did not match on volume; that is, participants loudly or quietly produced the sound /ba/ in their minds and listened to a quiet or loud recording of the sound /ba/, respectively—the dissimilar condition (see Figure 1I and L). The order of trials was random and different for each block, as well as different for each participant.

EEG Acquisition

We recorded the EEG with a BioSemi ActiveTwo system using 64 Ag/AgCl active electrodes placed according to the extended 10–20 system (FP1, FPz, FP2, AF7, AF3, AFz, AF4, AF8, F7, F5, F3, F1, Fz, F2, F4, F6, F8, FT7, FC5, FC3, FC1, FCz, FC2, FC4, FC6, FT8, T7, C5, C3, C1, Cz, C2, C4, C6, T8, TP7, CP5, CP3, CP1, CPz, CP2, CP4, CP6, TP8, P9, P7, P5, P3, P1, Pz, P2, P4, P6, P8, P10, PO7, PO3, POz, PO4, PO8, O1, Oz, O2, Iz). We also recorded the vertical EOG by placing an electrode above (we used FP1) and below the left eye, and the horizontal EOG by placing an electrode on the outer canthus of each eye. We also placed an electrode on the tip of the nose. The sampling rate of the EEG was 1024 Hz.

EEG Processing and ERP Analysis

We processed the EEG data using the protocol described in Chung and colleagues (2023) and Jack and colleagues (2019). We referenced the data to the electrode on the tip of the nose, and we filtered the data using a half-amplitude 0.5- to 30-Hz phase-shift free Butterworth filter (48 dB/Oct slope), as well as a 50-Hz Notch filter. We extracted the epochs from −2500 to 400 msec relative to audible sound onset, we corrected the epochs for eye-blink and movement artifacts using the technique described in Miller, Gratton, and Yee (1988) and Gratton, Coles, and Donchin (1983), and we excluded all epochs with signals exceeding peak-to-peak amplitudes of 200 μV at any EEG channel. We also excluded any epochs in which participants subsequently rated their performance on the trial as less than or equal to 3 out of 5. For the N1 analyses, we segmented the epochs from −100 to 400 msec, we baseline-corrected all epochs to their mean

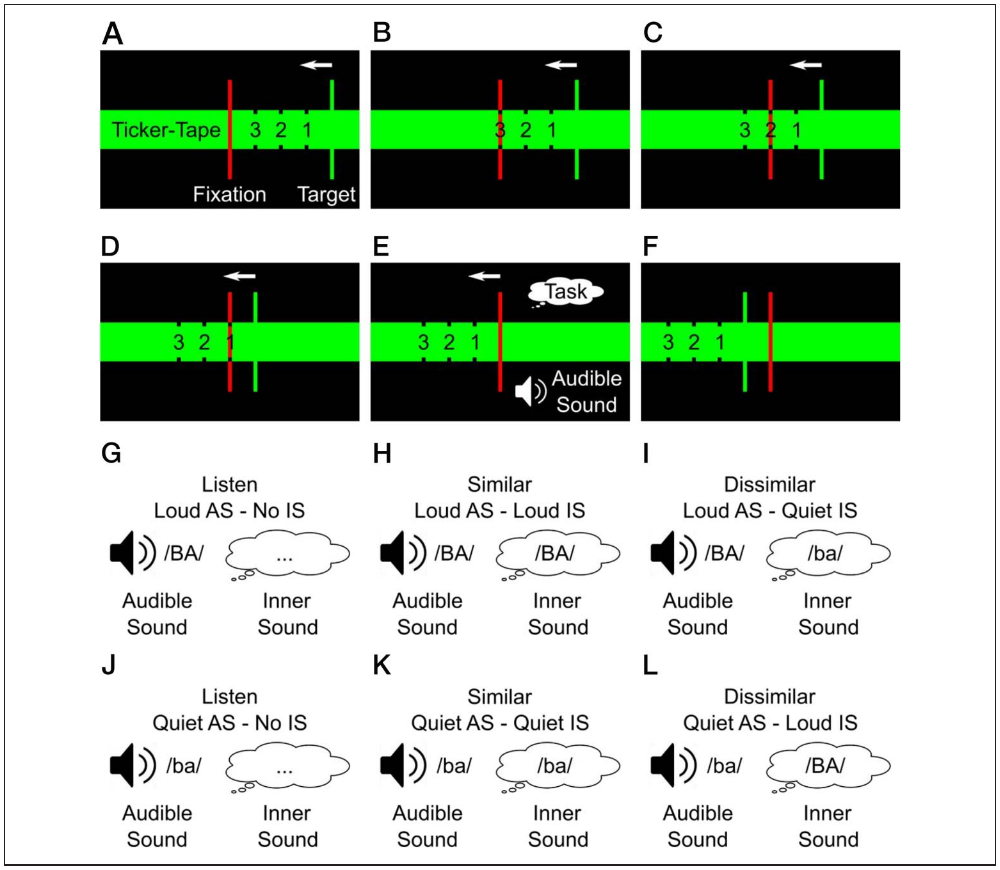

Figure 1. Schematic illustration of the animation and design. (A–F) Participants were instructed to look at the fixation line (which remained stationary) for the duration of the trial. After a short delay, the target line began to move leftward across the screen such that after 4 sec the target line overlapped the fixation line and subsequently continued to move across the ticker-tape. The audible sound /ba/ was presented at the precise moment that the target and fixation lines overlapped (as shown in E). (G, J) In the listen condition, participants passively listened to the audible sound. (H, I, K, L) In the inner speech conditions, participants silently produced the sound /ba/ in their minds. (H, L) On half of the trials in the inner speech condition, participants loudly produced the sound /ba/ in their minds—the loud inner sound condition; (I, K) on the other half of trials, participants quietly produced the sound /ba/ in their minds—the quiet inner sound condition. (H, K) Thus, on half of the trials, the inner and audible sounds matched on volume—the similar condition; (I, L) on the other half of trials, the inner and audible sounds did not match on volume—the dissimilar condition. Note that /BA/ is used to illustrate loud sounds whereas /ba/ is used to illustrate quiet sounds. AS = audible sound; IS = inner sound.

voltage from −100 to 0 msec, we computed an ERP for each condition, and we analyzed the mean amplitude of the ERPs averaged over Fz, FCz, and Cz electrodes in the time-window of 83–123 msec. We chose these electrodes to be consistent with Jack and colleagues (2019), and we selected this time-window using the collapsed localizer technique (Luck & Gaspelin, 2017). For the SNW analyses, we segmented the epochs from −2500 msec to 100 msec, we baseline-corrected all epochs to their mean voltage from −2500 to −2000 msec, we computed an ERP for each condition, and we analyzed the mean amplitude of the ERPs averaged over FCz, Cz, and CPz electrodes in the time-window of −500 to 0 msec. We chose these electrodes and this time-window to be consistent with Chung and colleagues (2023).

A limitation of frequentist analyses is that $p > .05$ can indicate either that the null hypothesis is supported or that there is a lack of evidence to distinguish between the null and alternative hypotheses. Therefore, we supplemented our analyses with Bayesian analyses to establish the extent to which the data provide support for the null or alternative hypothesis. The Bayes Factor (BF_{10}) is a continuous measure that is typically interpreted with respect to rule-of-thumb labels. For example, $BF_{10} < 1$ suggests anecdotal support for the null over the alternative hypothesis, $BF_{10} < 0.33$ suggests moderate support, and $BF_{10} < 0.1$ suggests strong support; whereas $BF_{10} > 1$ suggests anecdotal support for the alternative over the null hypothesis, $BF_{10} > 3$ suggests moderate support, and $BF_{10} > 10$ suggests strong support (Lee & Wagenmakers, 2013; Jeffreys, 1961). Bayesian analyses were conducted using the default Cauchy prior distribution ($r = .707$) in JASP, Version 0.16.3.0 (JASP Team, 2017).

RESULTS

Behavioral Results

Participants rated their subjective performance after each trial with a 5-point Likert scale, with scores ranging from 1, meaning *not at all successful*, to 5, meaning *completely successful*. Participants' mean ratings were 4.48 (SD = 0.46) in the listen condition, 4.36 (SD = 0.58) in the similar condition, and 4.49 (SD = 0.43) in the dissimilar condition: A one-way ANOVA with three levels (listen, similar, dissimilar) was significant, $F(2, 116) = 4.24$, $p = .017$, $\eta_p^2 = .07$, $BF_{10} = 2.05$. Post hoc comparisons found that the similar condition had significantly lower ratings than the listen condition, $t(58) = 2.04$, $p = .046$, $d = 0.27$, $BF_{10} = 0.98$, and that the similar condition had significantly lower ratings than the dissimilar condition, $t(58) = 2.81$, $p = .007$, $d = 0.37$, $BF_{10} = 4.98$, with no significant difference between the listen and dissimilar conditions, $t(58) = 0.29$, $p = .774$, $d = 0.04$, $BF_{10} = 0.15$. Moreover, participants' mean ratings were 4.43 (*SD* = 0.44) in the loud inner sound condition and 4.42 (*SD* = 0.56) in the quiet inner sound condition: A one-way ANOVA with three levels (listen, loud inner sound, and quiet inner sound) was not significant, $F(2, 116) = 0.87$, $p = .422$, $\eta_p^2 = .02$, $BF_{10} = 0.12$. This suggests that participants performed the task as instructed.

ERP Results

As discussed above, the aim of the present study was twofold. First, we sought to determine whether inner speech, similar to overt speech, is accompanied by a volume-dependent corollary discharge. To investigate this, we analyzed the mean amplitudes of the N1 and SNW with repeated-measures ANOVA using the factors Condition (listen, similar, dissimilar) and Volume (loud audible sound, quiet audible sound). Figure 2A shows the post-stimulus ERPs, Figure 2B shows the mean amplitudes for the N1 time-window, and Figure 2C shows the voltage maps for the N1 time-window. Analysis of the mean amplitudes for the N1 found a significant main effect of Condition, $F(2, 116) = 7.81$, $p < .001$, $\eta_p^2 = .12$, $BF_{10} = 8.98$. Post hoc comparisons found that the similar condition elicited a significantly smaller negative voltage than the listen condition, $t(58) = 2.74$, $p = .008$, $d = 0.36$, $BF_{10} = 4.24$, and that the dissimilar condition elicited a significantly smaller negative voltage than the listen condition, $t(58) = 3.28$, $p = .002$, $d = 0.43$, $BF_{10} = 16.28$, with no significant difference between the similar and dissimilar conditions, $t(58) = 0.43$, $p = .672$, $d = 0.06$, $BF_{10} = 0.16$ (see Figure 2D). There was also a significant main effect of Volume, $F(1, 58) = 4.33$, $p = .042$, $\eta_p^2 = .07$, $BF_{10} = 1.12$, indicating that loud audible sounds elicited a significantly larger negative voltage than quiet audible sounds. However, the interaction between Condition and Volume was

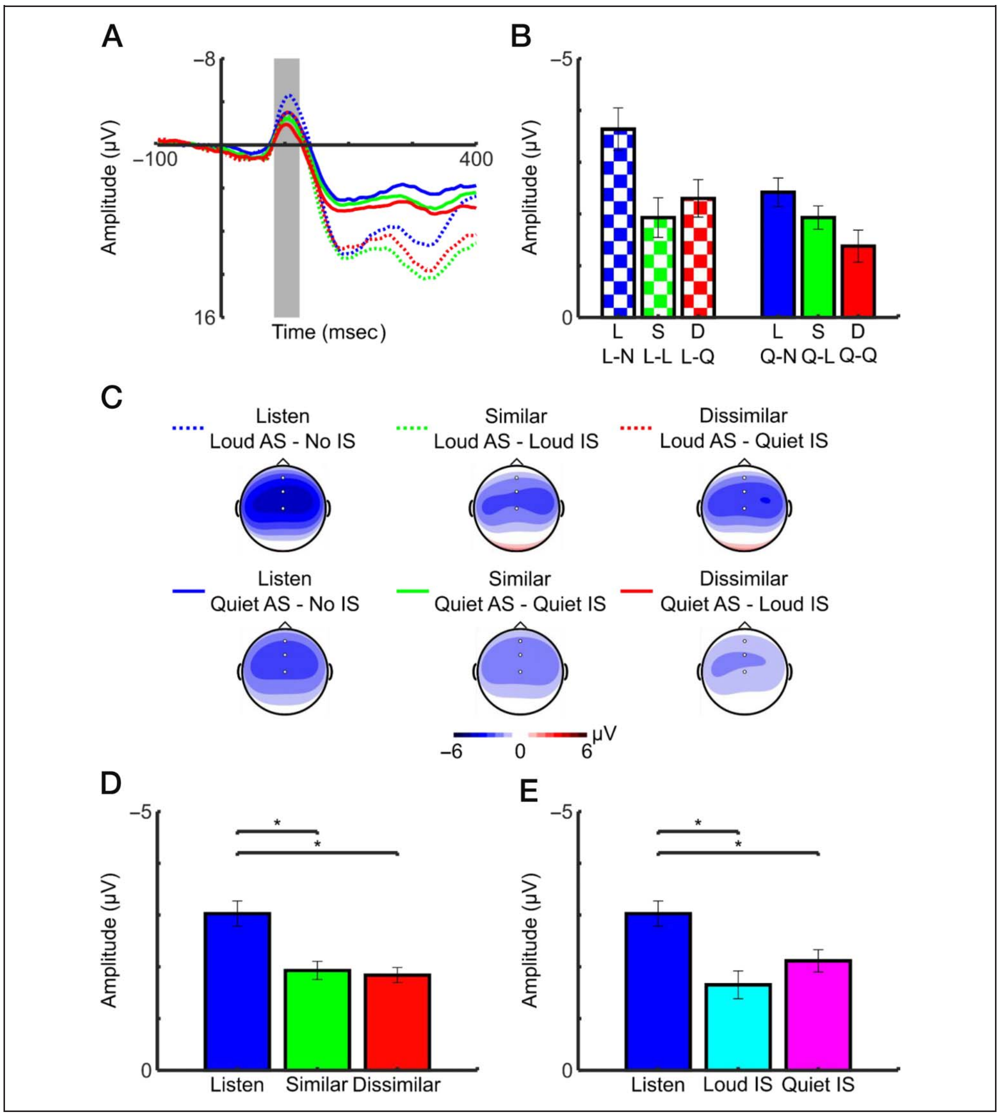

Figure 2. N1 results. (A) The line graph shows the grand-averaged ERPs for each condition averaged over Fz, FCz, and Cz electrodes, showing time (msec) on the *x* axis, with the placement of the *y* axis indicating the onset of the audible sound, and voltage (μV) on the *y* axis, with negative voltages plotted upward. The gray bar shows the N1 time-window (83–123 msec), which was selected using the collapsed localizer technique (Luck & Gaspelin, 2017). (B) The bar graph shows the mean amplitudes for each condition. (C) The voltage maps show the distribution of voltages over the scalp during the N1 time-window. (D) The bar graph shows the mean amplitudes for Aim 1: listen versus similar versus dissimilar. (E) The bar graph shows the mean amplitudes for Aim 2: listen versus loud inner sound versus quiet inner sound. Error bars show the *SEM*. L = Listen; S = Similar; D = Dissimilar; AS = audible sound; IS = inner sound; L-N = Loud AS-No IS; L-L = Loud AS-Loud IS; L-Q = Loud AS-Quiet IS; Q-N = Quiet AS-No IS; Q-L = Quiet AS-Loud IS; Q-Q = Quiet AS-Quiet IS.

not significant, $F(2, 116) = 1.49, p = .229, \eta_p^2 = .03$, $BF_{10} = 0.30$. These results show that producing the inner sound attenuated the N1 compared with listening, regardless of whether the volume of the inner sound was similar or dissimilar to the audible sound, with no difference between the similar and dissimilar conditions. This suggests that corollary discharges associated with inner speech do not contain information about the volume of inner speech.

Figure 3A shows the prestimulus ERPs, Figure 3B shows the mean amplitudes for the SNW time-window, and Figure 3C shows the voltage maps for the SNW time-window. Analysis of the mean amplitudes for the SNW found a significant main effect of Condition, $F(2, 116) = 7.85, p < .001, \eta_p^2 = .12$, $BF_{10} = 19.33$. Post hoc comparisons found that the similar condition elicited a significantly larger negative voltage than the listen condition, $t(58) = 3.60, p < .001, d = 0.47$, $BF_{10} = 39.27$, and that the dissimilar condition elicited a significantly larger negative voltage than the listen condition, $t(58) = 2.93, p = .005, d = 0.38$, $BF_{10} = 6.55$, with no significant difference between the similar and dissimilar conditions, $t(58) = 0.55, p = .583, d = 0.07$, $BF_{10} = 0.17$ (see Figure 3D). The main effect of Volume was not significant, $F(1, 58) = 0.01, p = .910, \eta_p^2 < .01$, $BF_{10} = 0.15$, indicating that there was no significant difference between loud and quiet audible sounds, nor was the interaction between Condition and Volume, $F(2, 116) = 1.72, p = .185, \eta_p^2 = .03$, $BF_{10} = 0.31$. These results show that producing the inner sound enhanced the SNW compared with listening, with no difference between the similar and dissimilar conditions. Similar to above, this suggests that corollary discharges are associated with inner speech.

The second aim of the present study was to determine whether the volume of inner speech influences the N1-attenuation effect or the magnitude of the SNW. To investigate this, we analyzed the mean amplitudes of the N1 and SNW and we reconfigured our repeated-measures ANOVA from listen versus similar versus dissimilar to listen versus loud inner sound versus quiet inner sound. Analysis of the mean amplitudes for the N1 found a

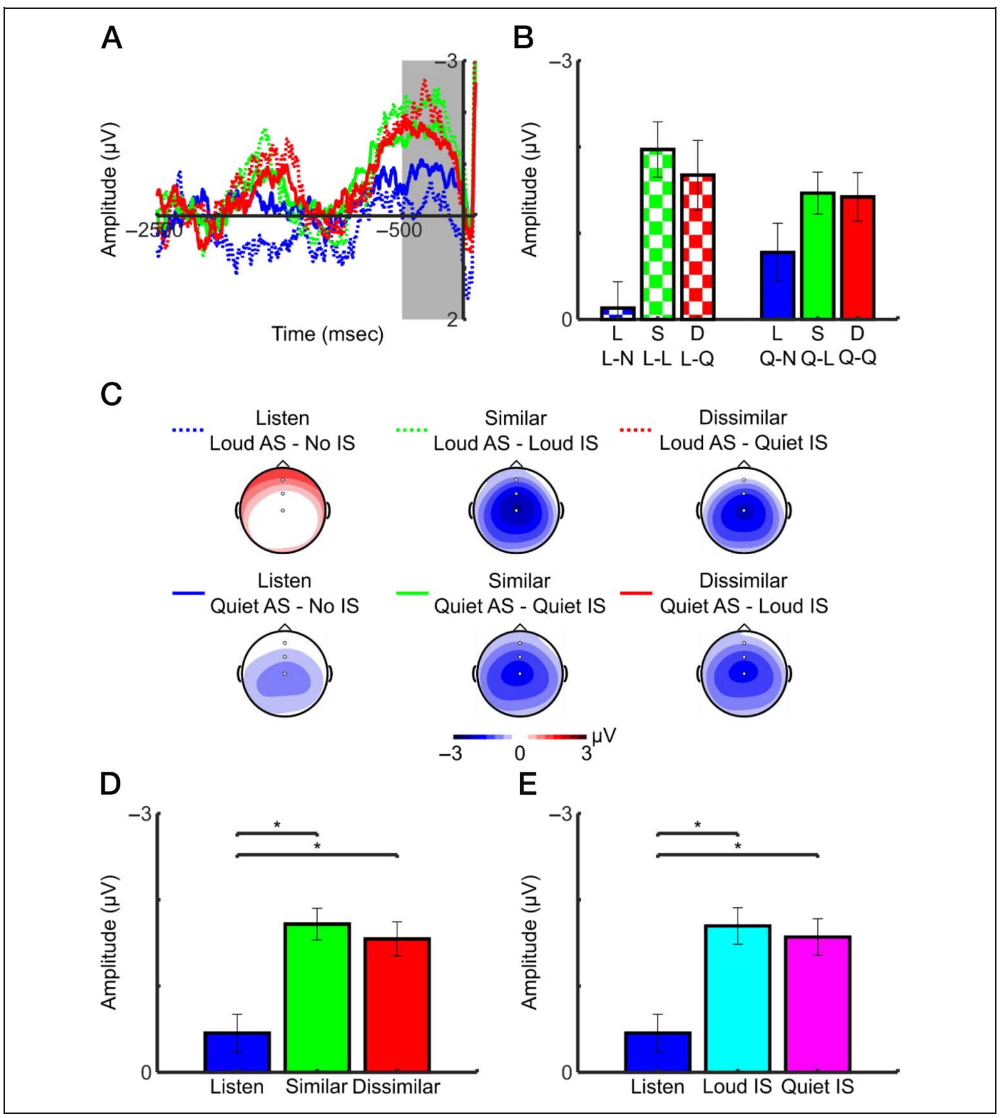

Figure 3. SNW results. (A) The line graph shows the grand-averaged ERPs for each condition averaged over Fz, FCz, and Cz electrodes, showing time (msec) on the *x* axis, with the placement of the *y* axis indicating the onset of the audible sound, and voltage (μV) on the *y* axis, with negative voltages plotted upward. The gray bar shows the SNW time-window (−500 to 0 msec), which was selected to be consistent with Chung and colleagues (2023). (B) The bar graph shows the mean amplitudes for each condition. (C) The voltage maps show the distribution of voltages over the scalp during the SNW time-window. (D) The bar graph shows the mean amplitudes for Aim 1: listen versus similar versus dissimilar. (E) The bar graph shows the mean amplitudes for Aim 2: listen versus loud inner sound versus quiet inner sound. Error bars show the *SEM*. L = Listen; S = Similar; D = Dissimilar; AS = audible sound; IS = inner sound; L-N = Loud AS-No IS; L-L = Loud AS-Loud IS; L-Q = Loud AS-Quiet IS; Q-N = Quiet AS-No IS; Q-L = Quiet AS-Loud IS; Q-Q = Quiet AS-Quiet IS.

significant main effect of Condition, $F(2, 116) = 5.48$, $p = .005$, $\eta_p^2 = .09$, $BF_{10} = 6.78$. Post hoc comparisons found that the loud inner sound condition elicited a significantly smaller negative voltage than the listen condition, $t(58) = 2.94$, $p = .005$, $d = 0.38$, $BF_{10} = 6.73$, and that the quiet inner sound condition elicited a significantly smaller negative voltage than the listen condition, $t(58) = 2.45$, $p = .017$, $d = 0.32$, $BF_{10} = 2.20$, with no significant difference between the loud and quiet inner sound conditions, $t(58) = 1.10$, $p = .277$, $d = 0.14$, $BF_{10} = 0.25$ (see Figure 2E). There was also a significant main effect of Volume, $F(1, 58) = 4.33$, $p = .042$, $\eta_p^2 = .07$, $BF_{10} = 1.12$, indicating that loud audible sounds elicited a significantly larger negative voltage than quiet audible sounds (note that this analysis is identical to the N1 analysis presented above, and is only repeated here for clarity). However, the interaction between Condition and Volume was not significant, $F(2, 116) = 1.44$, $p = .242$, $\eta_p^2 = .02$, $BF_{10} = 0.18$. These results show that producing the inner sound attenuated the N1 compared with listening, regardless of whether the volume of the inner sound was loud or quiet, with no difference between the loud and quiet inner sound conditions. This suggests that the volume of inner speech does not modulate the N1-attenuation effect.

Analysis of the mean amplitudes for the SNW found a significant main effect of Condition, $F(2, 116) = 6.82$, $p = .002$, $\eta_p^2 = .11$, $BF_{10} = 14.20$. Post hoc comparisons found that the loud inner sound condition elicited a significantly larger negative voltage than the listen condition, $t(58) = 3.32$, $p = .002$, $d = 0.43$, $BF_{10} = 18.10$, and that the quiet inner sound condition elicited a significantly larger negative voltage than the listen condition, $t(58) = 2.98$, $p = .004$, $d = 0.39$, $BF_{10} = 7.47$, with no significant difference between the loud and quiet inner sound conditions, $t(58) = 0.35$, $p = .727$, $d = 0.05$, $BF_{10} = 0.15$ (see Figure 3E). The main effect of Volume was not significant, $F(1, 58) = 0.01$, $p = .910$, $\eta_p^2 < .01$, $BF_{10} = 0.15$, indicating that there was no significant difference between loud and quiet audible sounds (note that this analysis is identical to the SNW analysis presented above, and is only repeated here for clarity), nor was the interaction between Condition and Volume, $F(2, 116) = 2.12$, $p = .125$, $\eta_p^2 = .04$, $BF_{10} = 0.41$. These results show that producing the inner sound enhanced the SNW compared with listening, with no difference between the loud and quiet inner sound conditions. Similar to above, this suggests that the volume of inner speech does not modulate the magnitude of the SNW.

DISCUSSION

The first aim of the present study was to determine whether inner speech, similar to overt speech (Bauer et al., 2006; Heinks-Maldonado & Houde, 2005), is accompanied by a volume-dependent corollary discharge. To investigate this, we used a similar procedure to Whitford and colleagues (2017), except that our participants were instructed to produce either a loud or quiet inner sound that was either similar or dissimilar to the volume of the audible sound. Assuming that corollary discharges contain information about the volume of inner speech, we hypothesized a larger N1-attenuation effect when the volume of the inner sound is similar to the audible sound compared with when it is not. Alternatively, if corollary discharges do not contain information about the volume of inner speech, we hypothesized similar N1-attenuation effects for the similar and dissimilar conditions. The results of the present study are consistent with the latter. That is, we found that producing the inner sound attenuated the N1 compared with listening, regardless of whether the volume of the inner sound was similar or dissimilar to the audible sound, with no difference between the similar and dissimilar conditions. We also found that producing the inner sound enhanced the SNW compared with listening, which is evidence of corollary discharge (Ody et al., 2023; Pinheiro et al., 2020; Reznik et al., 2018; Vercillo et al., 2018; Wen et al., 2018; Jo et al., 2014), and we replicated the well-known result that loud audible sounds elicit a larger N1 than quiet audible sounds (Mulert et al., 2005). This suggests that our failure to find a difference in N1-amplitude between the similar and dissimilar conditions was not because there was no corollary discharge, or because participants were unable to distinguish between loud and quiet audible sounds, or because participants' N1 was not differentially modulated by loud and quiet audible sounds; rather, it was because corollary discharges, despite containing information about the content and timing of inner speech (Jack et al., 2019; Whitford et al., 2017), might not contain information about the volume of inner speech. We conclude that there might be a functional difference between the neural processes that underlie the production of inner and overt speech, in that corollary discharges appear to contain information about the volume of overt speech (Bauer et al., 2006; Heinks-Maldonado & Houde, 2005) but not inner speech.

The key finding of the present study is that corollary discharges might not contain information about the volume of inner speech. We suspect that one reason for this is that volume does not serve a functional purpose in the context of inner speech because inner speech does not produce an audible sound. Specifically, one of the functions of corollary discharge is to protect our auditory receptors from desensitization caused by audible sounds (Crapse & Sommer, 2008). Evidence for this comes from monkeys and crickets, both of whom produce audible sounds as loud as 100 dB (Eliades & Wang, 2008; Poulet & Hedwig, 2002). However, because inner speech does not produce an audible sound, it is reasonable to assume that we are safe from desensitization during inner speech. One of the other functions of corollary discharge is to provide a sense of agency over self-generated sensory input

(Haggard, 2017; Gallagher, 2000), and it is possible that the content and timing of sensory input might be more important than the volume of sensory input for establishing a sense of agency. Anecdotal evidence for this comes from overt speech: If the content and timing of your overt speech is different to what you expected, then you are unlikely to have a sense of agency over those speech sounds (e.g., saying "dog" but hearing "cat," or saying "dog" but hearing "dog" several seconds later). In contrast, if the volume of your overt speech is different to what you expected, such as when your voice unintentionally echoes in a large concert hall, then you are likely to have a sense of agency over those speech sounds (e.g., saying "dog" but hearing "DOG"). It is possible that this might also apply to inner speech. Importantly, we are not suggesting that inner speech does not have volume; indeed, our participants and studies using self-report measures report that respondents can modulate the volume of their inner speech (Vilhauer, 2017; Cuevas-Yust, 2014). Instead, we are suggesting that volume does not serve a functional purpose in the context of inner speech because inner speech does not produce an audible sound. Nevertheless, because this has not been empirically tested, we concede that this is speculation.

If corollary discharges act to suppress the neural and perceptual responses to self-generated sensory input, and if inner speech does not produce an audible sound, then why is inner speech accompanied by a corollary discharge? One possibility, proposed by Pinker and Jackendoff (2005), is that inner speech evolved as a "by-product" of overt speech and thus continued to use many of the same neural processes, including corollary discharge. Another possibility, proposed by Jack and colleagues (2019), is that N1-attenuation elicited by inner speech is an example of the brain's prediction going too far, in that it generates an expectation of a sensory event that does not, and cannot, occur. Here, we propose a different possibility: Inner speech is a form of proprioception—one's ability to sense the position and action of their body in space (Sherrington, 1906)—for the mind, and proprioception is thought to rely on corollary discharge mechanisms (Latash, 2021; Skandalis, Lunsford, & Liao, 2021; Azim, Jiang, Alstermark, & Jessell, 2014). For instance, the act of waving your hand does not produce an audible sound, yet corollary discharges are constantly predicting the next position, or state, of your hand as it moves back and forward through space, allowing us to know its precise location. Similarly, even though inner speech does not produce an audible sound, it is possible that corollary discharges predict the next state of inner speech, the output of which is our inner voice, which we are conscious of and have a sense of agency over. This hypothesis aligns with the popular notion that inner speech is an action without motor output (Jones & Fernyhough, 2007; Frith, 1987; Feinberg, 1978) and is similar to recent attempts to apply the internal forward model to cognitive domains, such as decision-making (Subramanian, Alers, & Sommer, 2019). Therefore, we propose that the N1-attenuation effects observed in the present study might reflect the extent to which the brain is making a prediction about the state of inner speech. We look forward to exploring this hypothesis in the future.

So far, we have interpreted the absence of a difference in N1-amplitude between the similar and dissimilar conditions to mean that corollary discharges might not contain information about the volume of inner speech. However, an alternative explanation for the absence of a difference is that participants did not produce loud and quiet inner sounds as instructed. This is a concern for the field (and for many areas of psychological science) because inner speech is a private experience with no behavioral manifestation; thus, it is difficult to know whether participants performed the task or not. Indeed, the overarching aim of the field is to develop an objective measure to overcome this limitation. Nevertheless, we think this alternative explanation is unlikely for at least three reasons. First, our participants acknowledged that they could produce loud and quiet inner sounds during the informed consent process, and they were able to practice producing loud and quiet inner sounds before the experiment. Second, our participants rated their ability to produce loud and quiet inner sounds after every trial, and our participants' mean ratings were approximately the same as those reported by Jack and colleagues (2019) and Whitford and colleagues (2017), both of whom found a difference in N1-amplitude between the match and mismatch conditions. Third, our participants did something in the inner speech blocks, which they did not do in the listen blocks, as evidenced by the difference in N1- and SNW-amplitude between the listen and inner speech conditions. However, because the order of similar and dissimilar trials was random and different for each block and for each participant, it is unlikely that participants had foreknowledge about whether they would hear a loud or quiet audible sound on any given trial and thus modulated the volume of the inner sound across trials accordingly. Taken together, this suggests that the results of the present study are not because participants did not perform the task as instructed; rather, it is because corollary discharges might not contain information about the volume of inner speech.

The second aim of the present study was to determine whether the volume of inner speech influences the N1-attenuation effect or the magnitude of the SNW. To investigate this, we reconfigured our repeated-measures ANOVA from listen versus similar versus dissimilar to listen versus loud inner sound versus quiet inner sound. We found that producing the inner sound attenuated the N1 and enhanced the SNW compared with listening, regardless of whether the volume of the inner sound was loud or quiet, with no difference between the loud and quiet inner sound conditions. Even though we found that the difference in N1-amplitude between the loud and quiet inner sound conditions was not significant (p = .277), the mean values suggest a trend toward a larger N1-attenuation effect for loud inner sounds ($M = -1.38\ \mu V$) than for quiet inner sounds ($M = -0.91\ \mu V$), a result broadly consistent with Tian and colleagues (2018), who reported a significant difference in N1-attenuation between these conditions (they did not analyze the SNW). One explanation for the lack of significance in the present study is that the production of inner speech might require multiple cognitive mechanisms (Pratts, Pobric, & Yao, 2023; Ma & Tian, 2019; Alderson-Day, Mitrenga, Wilkinson, McCarthy-Jones, & Fernyhough, 2018; Hurlburt, Alderson-Day, Kühn, & Fernyhough, 2016; Tian et al., 2016), and it is possible that our procedure isolated a different mechanism from Tian and colleagues' (2018) procedure. Specifically, one difference is that our participants produced the loud or quiet inner sound at the precise moment they heard the audible sound whose volume randomly varied between trials, whereas their participants produced the loud or quiet inner sound four times before hearing an audible sound whose volume remained constant across trials. Another possibility is that our "similar versus dissimilar" contrast interfered with detecting the "loud versus quiet" effect reported by Tian and colleagues (2018). Another difference is that we measured N1 using conventional ERP analyses, which average voltage at individual electrodes to examine brain activity at specific spatial and temporal regions of interest (Luck, 2014), whereas they used global field power, which calculates the standard deviation of voltage at all electrodes to assess the overall strength of brain activity (Lehmann & Skrandies, 1980; Lehmann, 1971). We suspect that these procedural differences might have reduced the sensitivity of the present study to detect whether the volume of inner speech influences the N1-attenuation effect or the magnitude of the SNW.

We would now like to comment on the functional significance of the SNW. As mentioned above, we found that producing the inner sound enhanced the SNW compared with listening. This replicates the results of Chung and colleagues (2023) and is consistent with the idea that the SNW is evidence of corollary discharge (Ody et al., 2023; Pinheiro et al., 2020; Reznik et al., 2018; Vercillo et al., 2018; Wen et al., 2018; Jo et al., 2014). However, it is important to acknowledge that the SNW might reflect several different processes. Depending on the procedure, the SNW could be a readiness potential (RP; or "Bereitschaftspotential"), which is thought to index motor preparation for a voluntary action (Schurger et al., 2021; Brunia et al., 2012; Kornhuber & Deecke, 1965), a stimulus preceding negative (SPN), which is thought to index the expectation of an upcoming stimulus that provides important information about the past or future (Brunia et al., 2012; van Boxtel & Böcker, 2004; Damen & Brunia, 1987), or a contingent negative variation (CNV), which is thought to index the expectation of an upcoming stimulus that requires a motor response (Brunia et al., 2012; van Boxtel & Böcker, 2004; Walter, Cooper, Aldridge, McCallum, & Winter, 1964). Given that the animation instructs participants about when to produce the inner sound as well as when they will hear the audible sound, Chung and colleagues (2023) concluded that their SNW is likely to be the CNV; however, they conceded it could also be the RP, which has been found for other imagined actions (Park et al., 2022; Pinheiro et al., 2020; Niazi et al., 2013; do Nascimento et al., 2006; Jankelowitz & Colebatch, 2002; Cunnington, Iansek, Johnson, & Bradshaw, 1997; Cunnington, Iansek, Bradshaw, & Phillips, 1996; Beisteiner, Höllinger, Lindinger, Lang, & Berthoz, 1995). It is also possible that the SNW is the SPN elicited by the expectation of the upcoming audible sound, or some combination of the RP and SPN. Indeed, it has been proposed that the CNV might actually be a combination of the RP and SPN (Damen & Brunia, 1987; but see Kotani, Ohgami, Arai, Kiryu, & Inoue, 2011). Unfortunately, the ambiguous relationship between RP, SPN, and CNV makes identifying the functional significance of the SNW difficult. We look forward to distinguishing between these components in the future.

Among the limitations of the present study is the fact that the listen, loud inner speech, and quiet inner speech conditions were presented in different blocks of trials. This raises the possibility that differences in task context or attention might have contributed to the observed N1-attenuation effects. While previous research suggests that N1-attenuation is robust to variations in task context (Baess, Horváth, Jacobsen, & Schröger, 2011) and attention (Saupe, Widmann, Trujillo-Barreto, & Schröger, 2013; Timm, SanMiguel, Saupe, & Schröger, 2013), the block-wise separation of the listen, loud inner speech, and quiet inner speech conditions nonetheless represents a potential confound. Notably, this limitation does not affect comparisons made between the similar and dissimilar conditions because they occurred within the same blocks of trials. Another potential limitation is that we used a standardized recording of the audible sound for every participant rather than a recording of each participant's own overt voice. On the assumption that a participants' inner voice will typically resemble their overt voice (Filik & Barber, 2011; but see Oppenheim & Dell, 2008), it is possible that using a recording of participants' own overt voice as the audible sound might lead to a larger N1-attenuation effect compared with a standardized voice. However, this approach introduces certain challenges. In particular, using a different audible sound for each participant can result in larger between-participant variability in both N1-amplitude and latency, especially when including participants whose overt voices may differ in their auditory properties. In contrast, using a standardized voice allows for greater control over the auditory properties of the audible sound, thereby reducing variability. Moreover, use of a standardized voice aligns the present study with previous research employing the same animation and general procedure (Whitford et al., 2017, 2025; Chung et al., 2023; Jack et al., 2019), ensuring consistency and facilitating direct comparisons with prior findings. To the best of our knowledge, there are no studies directly comparing N1-attenuation using a standardized voice versus participants' own overt voice as the audible sound. We look forward to conducting this experiment in the future.

Finally, although the aim of the present study was to understand the neural processes underlying inner speech, our results have important implications for our understanding of dysfunctions of inner speech, such as auditory–verbal hallucinations, which are highly characteristic of schizophrenia (Fletcher & Frith, 2009; Mellor, 1970). From a cognitive standpoint, auditory–verbal hallucinations are thought to be the misattribution of inner speech to an external source (Frith, 1987; Feinberg, 1978), presumably because the operation of the internal forward model is impaired, such as by the delayed transmission of corollary discharges (Whitford, Ford, Mathalon, Kubicki, & Shenton, 2012; Whitford et al., 2011). Consistent with this, patients who experience auditory–verbal hallucinations typically describe their hallucinations and thoughts as being the same volume as their overt speech, whereas patients who do not experience auditory–verbal hallucinations typically describe their overt speech as being louder than their thoughts (Cuevas-Yust, 2014). This suggests that turning down the volume of auditory–verbal hallucinations might lead to significant improvements in the severity of these symptoms. However, even if cognitive training or pharmaceutical treatments were able to accomplish this goal, these remedies might have little or no effect on the underlying neural processes that give rise to auditory–verbal hallucinations because corollary discharges might not contain information about the volume of inner speech. Furthermore, the results of the present study might be useful for the ongoing development of brain–computer interfaces aimed at deciphering inner speech for people who are unable to produce overt speech (Lebedev & Nicolelis, 2006). Specifically, even though N1-attenuation can be used to identify the content and timing of inner speech (Jack et al., 2019; Whitford et al., 2017), the results of the present study suggest that N1-attenuation might not reflect the volume of inner speech. Therefore, the goal of future research should be to identify brain activity correlating with the volume of inner speech.

Corresponding author: Bradley N. Jack, Research School of Psychology, Australian National University, Canberra, Australia, e-mail: bradley.jack@anu.edu.au.

Data Availability Statement

Data for this study are available upon request to the corresponding author.

Author Contributions

Kevin Berryman: Conceptualization; Formal analysis; Investigation; Methodology; Visualization; Writing—Original draft; Writing—Review & editing. Thomas J. Whitford: Conceptualization; Writing—Review & editing. Mike E. Le Pelley: Conceptualization; Writing—Review & editing. Bradley N. Jack: Conceptualization; Data curation; Formal analysis; Funding acquisition; Methodology; Project administration; Resources; Software; Supervision; Validation; Visualization; Writing—Original draft; Writing—Review & editing.

Funding Information

This work was supported by the Australian Research Council (https://dx.doi.org/10.13039/501100000923), grant number: DE220100739.

Diversity in Citation Practices

Retrospective analysis of the citations in every article published in this journal from 2010 to 2021 reveals a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience* (*JoCN*) during this period were M(an)/M = .407, W(oman)/M = .32, M/W = .115, and W/W = .159, the comparable proportions for the articles that these authorship teams cited were M/M = .549, W/M = .257, M/W = .109, and W/W = .085 (Postle and Fulvio, *JoCN*, 34:1, pp. 1–3). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article's gender citation balance.

REFERENCES

Alderson-Day, B., & Fernyhough, C. (2015). Inner speech: Development, cognitive functions, phenomenology, and neurobiology. *Psychological Bulletin*, *141*, 931–965. https://doi.org/10.1037/bul0000021, PubMed: 26011789

Alderson-Day, B., Mitrenga, K., Wilkinson, S., McCarthy-Jones, S., & Fernyhough, C. (2018). The Varieties of Inner Speech Questionnaire–Revised (VISQ-R): Replicating and refining links between inner speech and psychopathology. *Consciousness and Cognition*, *65*, 48–58. https://doi.org/10.1016/j.concog.2018.07.001, PubMed: 30041067

Azim, E., Jiang, J., Alstermark, B., & Jessell, T. M. (2014). Skilled reaching relies on a V2a propriospinal internal copy circuit. *Nature*, *508*, 357–363. https://doi.org/10.1038/nature13021, PubMed: 24487617

Baess, P., Horváth, J., Jacobsen, T., & Schröger, E. (2011). Selective suppression of self-initiated sounds in an auditory stream: An ERP study. *Psychophysiology*, *48*, 1276–1283. https://doi.org/10.1111/j.1469-8986.2011.01196.x, PubMed: 21449953

Bauer, J. J., Mittal, J., Larson, C. R., & Hain, T. J. (2006). Vocal responses to unanticipated perturbations in voice loudness feedback: An automatic mechanism for stabilizing voice amplitude. *Journal of the Acoustical Society of America*, *119*, 2363–2371. https://doi.org/10.1121/1.2173513, PubMed: 16642849

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*, 1303–1312. https://doi.org/10.1016/j.clinph.2009.04.022, PubMed: 19520602
Behroozmand, R., & Larson, C. R. (2011). Error-dependent modulation of speech-induced auditory suppression for pitch-shifted voice feedback. *BMC Neuroscience*, *12*, 54. https://doi.org/10.1186/1471-2202-12-54, PubMed: 21645406
Behroozmand, R., Liu, H., & Larson, C. R. (2011). Time-dependent neural processing of auditory feedback during voice pitch error detection. *Journal of Cognitive Neuroscience*, *23*, 1205–1217. https://doi.org/10.1162/jocn.2010.21447, PubMed: 20146608
Behroozmand, R., Sangtian, S., Korzyukov, O., & Larson, C. R. (2016). A temporal predictive code for voice motor control: Evidence from ERP and behavioral responses to pitch-shifted auditory feedback. *Brain Research*, *1636*, 1–12. https://doi.org/10.1016/j.brainres.2016.01.040, PubMed: 26835556
Beisteiner, R., Höllinger, P., Lindinger, G., Lang, W., & Berthoz, A. (1995). Mental representations of movements. Brain potentials associated with imagination of hand movements. *Electroencephalography and Clinical Neurophysiology*, *96*, 183–193. https://doi.org/10.1016/0168-5597(94)00226-5, PubMed: 7535223
Bendixen, A., SanMiguel, I., & Schröger, E. (2012). Early electrophysiological indicators for predictive processing in audition: A review. *International Journal of Psychophysiology*, *83*, 120–131. https://doi.org/10.1016/j.ijpsycho.2011.08.003, PubMed: 21867734
Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436. PubMed: 9176952
Brunia, C. H. M., van Boxtel, G. J. M., & Böcker, K. B. E. (2012). Negative slow waves as indices of anticipation: The bereitschaftspotential, the contingent negative variation, and the stimulus-preceding negativity. In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford handbook of event-related potential components* (pp. 190–208). Oxford, UK: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0108
Chen, Z., Chen, X., Liu, P., Huang, D., & Liu, H. (2012). Effect of temporal predictability on the neural processing of self-triggered auditory stimulation during vocalization. *BMC Neuroscience*, *13*, 55. https://doi.org/10.1186/1471-2202-13-55, PubMed: 22646514
Chung, L. K.-H., Jack, B. N., Griffiths, O., Pearson, D., Luque, D., Harris, A. W. F., et al. (2023). Neurophysiological evidence of motor preparation in inner speech and the effect of content predictability. *Cerebral Cortex*, *33*, 11556–11569. https://doi.org/10.1093/cercor/bhad389, PubMed: 37943760
Crapse, T. B., & Sommer, M. A. (2008). Corollary discharge circuits in the primate brain. *Current Opinion in Neurobiology*, *18*, 552–557. https://doi.org/10.1016/j.conb.2008.09.017, PubMed: 18848626
Cuevas-Yust, C. (2014). Do thoughts have sound? Differences between thoughts and auditory hallucinations in schizophrenia. *Spanish Journal of Psychology*, *17*, E27. https://doi.org/10.1017/sjp.2014.29, PubMed: 25011699
Cunnington, R., Iansek, R., Bradshaw, J. L., & Phillips, J. G. (1996). Movement-related potentials associated with movement preparation and motor imagery. *Experimental Brain Research*, *111*, 429–436. https://doi.org/10.1007/BF00228732, PubMed: 8911937
Cunnington, R., Iansek, R., Johnson, K. A., & Bradshaw, J. L. (1997). Movement-related potentials in Parkinson's disease. Motor imagery and movement preparation. *Brain*, *120*, 1339–1353. https://doi.org/10.1093/brain/120.8.1339, PubMed: 9278627
Damen, E. J. P., & Brunia, C. H. M. (1987). Changes in heart rate and slow brain potentials related to motor preparation and stimulus anticipation in a time estimation task. *Psychophysiology*, *24*, 700–713. https://doi.org/10.1111/j.1469-8986.1987.tb00353.x, PubMed: 3438435
do Nascimento, O. F., Nielsen, K. D., & Voigt, M. (2006). Movement-related parameters modulate cortical activity during imaginary isometric plantar-flexions. *Experimental Brain Research*, *171*, 78–90. https://doi.org/10.1007/s00221-005-0247-z, PubMed: 16320044
Eliades, S. J., & Wang, X. (2008). Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature*, *453*, 1102–1106. https://doi.org/10.1038/nature06910, PubMed: 18454135
Feinberg, I. (1978). Efference copy and corollary discharge: Implications for thinking and its disorders. *Schizophrenia Bulletin*, *4*, 636–640. https://doi.org/10.1093/schbul/4.4.636, PubMed: 734369
Filik, R., & Barber, E. (2011). Inner speech during silent reading reflects the reader's regional accent. *PLoS One*, *6*, e25782. https://doi.org/10.1371/journal.pone.0025782, PubMed: 22039423
Fletcher, P. C., & Frith, C. D. (2009). Perceiving is believing: A Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews Neuroscience*, *10*, 48–58. https://doi.org/10.1038/nrn2536, PubMed: 19050712
Ford, J. M., & Mathalon, D. H. (2004). Electrophysiological evidence of corollary discharge dysfunction in schizophrenia during talking and thinking. *Journal of Psychiatric Research*, *38*, 37–46. https://doi.org/10.1016/S0022-3956(03)00095-5, PubMed: 14690769
Frith, C. D. (1987). The positive and negative symptoms of schizophrenia reflect impairments in the perception and initiation of action. *Psychological Medicine*, *17*, 631–648. https://doi.org/10.1017/S0033291700025873, PubMed: 3628624
Gallagher, S. (2000). Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Sciences*, *4*, 14–21. https://doi.org/10.1016/S1364-6613(99)01417-5, PubMed: 10637618
Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*, 468–484. https://doi.org/10.1016/0013-4694(83)90135-9, PubMed: 6187540
Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, *18*, 196–207. https://doi.org/10.1038/nrn.2017.14, PubMed: 28251993
Hall, L., Dawel, A., Greenwood, L.-M., Monaghan, C., Berryman, K., & Jack, B. N. (2023). Estimating statistical power for ERP studies using the auditory N1, Tb, and P2 components. *Psychophysiology*, *60*, e14363. https://doi.org/10.1111/psyp.14363, PubMed: 37382363
Heinks-Maldonado, T. H., & Houde, J. F. (2005). Compensatory responses to brief perturbations of speech amplitude. *Acoustic Research Letters Online*, *6*, 131–137. https://doi.org/10.1121/1.1931747
Heinks-Maldonado, T. H., Mathalon, D. H., Gray, M., & Ford, J. M. (2005). Fine-tuning of auditory cortex during speech production. *Psychophysiology*, *42*, 180–190. https://doi.org/10.1111/j.1469-8986.2005.00272.x, PubMed: 15787855
Horváth, J. (2015). Action-related auditory ERP attenuation: Paradigms and hypotheses. *Brain Research*, *1626*, 54–65. https://doi.org/10.1016/j.brainres.2015.03.038, PubMed: 25843932
Houde, J. F., Nagarajan, S. S., Sekihara, K., & Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: An MEG study. *Journal of Cognitive Neuroscience*, *14*, 1125–1138. https://doi.org/10.1162/089892902760807140, PubMed: 12495520

Hughes, G., Desantis, A., & Waszak, F. (2013). Mechanisms of intentional binding and sensory attenuation: The role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychological Bulletin*, *139*, 133–151. https://doi.org/10.1037/a0028566, PubMed: 22612280

Hurlburt, R. T., Alderson-Day, B., Kühn, S., & Fernyhough, C. (2016). Exploring the ecological validity of thinking on demand: Neural correlates of elicited vs. spontaneously occurring inner speech. *PLoS One*, *11*, e0147932. https://doi.org/10.1371/journal.pone.0147932, PubMed: 26845028

Jack, B. N., Le Pelley, M. E., Han, N., Harris, A. W. F., Spencer, K. M., & Whitford, T. J. (2019). Inner speech is accompanied by a temporally-precise and content-specific corollary discharge. *Neuroimage*, *198*, 170–180. https://doi.org/10.1016/j.neuroimage.2019.04.038, PubMed: 31002966

Jankelowitz, S. K., & Colebatch, J. G. (2002). Movement-related potentials associated with self-paced, cued and imagined arm movements. *Experimental Brain Research*, *147*, 98–107. https://doi.org/10.1007/s00221-002-1220-8, PubMed: 12373374

JASP Team. (2017). *JASP* (Version 0.16.3.0) [Computer software].

Jeffreys, H. (1961). *The theory of probability*. Oxford, UK: Clarendon Press.

Jo, H.-G., Wittmann, M., Hinterberger, T., & Schmidt, S. (2014). The readiness potential reflects intentional binding. *Frontiers in Human Neuroscience*, *8*, 421. https://doi.org/10.3389/fnhum.2014.00421, PubMed: 24959135

Jones, S. R., & Fernyhough, C. (2007). Thought as action: Inner speech, self-monitoring, and auditory verbal hallucinations. *Consciousness and Cognition*, *16*, 391–399. https://doi.org/10.1016/j.concog.2005.12.003, PubMed: 16464616

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception*, *36*, 14.

Kornhuber, H. H., & Deecke, L. (1965). Changes in the brain potential in voluntary movements and passive movements in man: Readiness potential and reafferent potentials. *Pflügers Archiv für die Gesamte Physiologie des Menschen und der Tiere*, *284*, 1–17. PubMed: 14341490

Kotani, Y., Ohgami, Y., Arai, J., Kiryu, S., & Inoue, Y. (2011). Motor and nonmotor components of event-brain potential in preparation of motor response. *Journal of Behavioral and Brain Science*, *1*, 234–241. https://doi.org/10.4236/jbbs.2011.14030

Latash, M. L. (2021). Laws of nature that define biological action and perception. *Physics of Life Reviews*, *36*, 47–67. https://doi.org/10.1016/j.plrev.2020.07.007, PubMed: 32868159

Lebedev, M. A., & Nicolelis, M. A. L. (2006). Brain–machine interfaces: Past, present and future. *Trends in Neurosciences*, *29*, 536–546. https://doi.org/10.1016/j.tins.2006.07.004, PubMed: 16859758

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modelling: A practical course*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139087759

Lehmann, D. (1971). Multichannel topography of human alpha EEG fields. *Electroencephalography and Clinical Neurophysiology*, *31*, 439–449. https://doi.org/10.1016/0013-4694(71)90165-9, PubMed: 4107798

Lehmann, D., & Skrandies, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalography and Clinical Neurophysiology*, *48*, 609–621. https://doi.org/10.1016/0013-4694(80)90419-8, PubMed: 6155251

Liu, H., Meshman, M., Behroozmand, R., & Larson, C. R. (2011). Differential effects of perturbation direction and magnitude on the neural processing of voice pitch feedback. *Clinical Neurophysiology*, *122*, 951–957. https://doi.org/10.1016/j.clinph.2010.08.010, PubMed: 20869305

Luck, S. J. (2014). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology*, *54*, 146–157. https://doi.org/10.1111/psyp.12639, PubMed: 28000253

Ma, O., & Tian, X. (2019). Distinct mechanisms of imagery differentially influence speech perception. *eNeuro*, *6*, ENEURO.0261-19.2019. https://doi.org/10.1523/ENEURO.0261-19.2019, PubMed: 31481396

Mellor, C. S. (1970). First rank symptoms of schizophrenia: I. The frequency in schizophrenics on admission to hospital II. Differences between individual first rank symptoms. *British Journal of Psychiatry*, *117*, 15–23. https://doi.org/10.1192/S0007125000192116

Miller, G. A., Gratton, G., & Yee, C. M. (1988). Generalized implementation of an eye movement correction procedure. *Psychophysiology*, *25*, 241–243. https://doi.org/10.1111/j.1469-8986.1988.tb00999.x

Mulert, C., Jäger, L., Propp, S., Karch, S., Störmann, S., Pogarell, O., et al. (2005). Sound level dependence of the primary auditory cortex: Simultaneous measurement with 61-channel EEG and fMRI. *Neuroimage*, *28*, 49–58. https://doi.org/10.1016/j.neuroimage.2005.05.041, PubMed: 16006148

Näätänen, R., & Picton, T. (1987). 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. https://doi.org/10.1111/j.1469-8986.1987.tb00311.x, PubMed: 3615753

Niazi, I. K., Jiang, N., Jochumsen, M., Nielsen, J. F., Dremstrup, K., & Farina, D. (2013). Detection of movement-related cortical potentials based on subject-independent training. *Medical & Biological Engineering & Computing*, *51*, 507–512. https://doi.org/10.1007/s11517-012-1018-1, PubMed: 23283643

Ody, E., Kircher, T., Straube, B., & He, Y. (2023). Pre-movement event-related potentials and multivariate pattern of EEG encode action outcome prediction. *Human Brain Mapping*, *44*, 6198–6213. https://doi.org/10.1002/hbm.26506, PubMed: 37792296

Oppenheim, G. M., & Dell, G. S. (2008). Inner speech slips exhibit lexical bias, but not the phonemic similarity effect. *Cognition*, *106*, 528–537. https://doi.org/10.1016/j.cognition.2007.02.006, PubMed: 17407776

Park, H.-D., Piton, T., Kannape, O. A., Duncan, N. W., Lee, K.-Y., Lane, T. J., et al. (2022). Breathing is coupled with voluntary initiation of mental imagery. *Neuroimage*, *264*, 119685. https://doi.org/10.1016/j.neuroimage.2022.119685, PubMed: 36252914

Pelli, D. G. (1997). The Videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442. PubMed: 9176953

Perrone-Bertolotti, M., Rapin, L., Lachaux, J. P., Baciu, M., & Lœvenbruck, H. (2014). What is that little voice inside my head? Inner speech phenomenology, its role in cognitive performance, and its relation to self-monitoring. *Behavioural Brain Research*, *261*, 220–239. https://doi.org/10.1016/j.bbr.2013.12.034, PubMed: 24412278

Pinheiro, A. P., Schwartze, M., Gutiérrez-Domínguez, F., & Kotz, S. A. (2020). Real and imagined sensory feedback have comparable effects on action anticipation. *Cortex*, *130*, 290–301. https://doi.org/10.1016/j.cortex.2020.04.030, PubMed: 32698087

Pinker, S., & Jackendoff, R. (2005). The faculty of language: What's special about it? *Cognition*, *95*, 201–236. https://doi.org/10.1016/j.cognition.2004.08.004, PubMed: 15694646

Poulet, J. F. A., & Hedwig, B. (2002). A corollary discharge maintains auditory sensitivity during sound production.

Nature, *418*, 872–876. https://doi.org/10.1038/nature00919, PubMed: 12192409
Poulet, J. F. A., & Hedwig, B. (2007). New insights into corollary discharges mediated by identified neural pathways. *Trends in Neurosciences*, *30*, 14–21. https://doi.org/10.1016/j.tins.2006.11.005, PubMed: 17137642
Pratts, J., Pobric, G., & Yao, B. (2023). Bridging phenomenology and neural mechanisms of inner speech: ALE meta-analysis on egocentricity and spontaneity in a dual-mechanistic framework. *Neuroimage*, *282*, 120399. https://doi.org/10.1016/j.neuroimage.2023.120399, PubMed: 37827205
Reznik, D., Simon, S., & Mukamel, R. (2018). Predicted sensory consequences of voluntary actions modulate amplitude of preceding readiness potentials. *Neuropsychologia*, *119*, 302–307. https://doi.org/10.1016/j.neuropsychologia.2018.08.028, PubMed: 30172828
Saupe, K., Widmann, A., Trujillo-Barreto, N. J., & Schröger, E. (2013). Sensorial suppression of self-generated sounds and its dependence on attention. *International Journal of Psychophysiology*, *90*, 300–310. https://doi.org/10.1016/j.ijpsycho.2013.09.006, PubMed: 24095710
Schröger, E., Marzecová, A., & SanMiguel, I. (2015). Attention and prediction in human audition: A lesson from cognitive psychophysiology. *European Journal of Neuroscience*, *41*, 641–664. https://doi.org/10.1111/ejn.12816, PubMed: 25728182
Schurger, A., Hu, P., Pak, J., & Roskies, A. L. (2021). What is the readiness potential? *Trends in Cognitive Sciences*, *25*, 558–570. https://doi.org/10.1016/j.tics.2021.04.001, PubMed: 33931306
Schütz-Bosbach, S., & Prinz, W. (2007). Perceptual resonance: Action-induced modulation of perception. *Trends in Cognitive Sciences*, *11*, 349–355. https://doi.org/10.1016/j.tics.2007.06.005, PubMed: 17629544
Scott, M. (2013). Corollary discharge provides the sensory content of inner speech. *Psychological Science*, *24*, 1824–1830. https://doi.org/10.1177/0956797613478614, PubMed: 23846719
Sherrington, C. S. (1906). *The integrative action of the nervous system*. New Haven: Yale University Press.
Sitek, K. R., Mathalon, D. H., Roach, B. J., Houde, J. F., Niziolek, C. A., & Ford, J. M. (2013). Auditory cortex processes variation in our own speech. *PLoS One*, *8*, e82925. https://doi.org/10.1371/journal.pone.0082925, PubMed: 24349399
Skandalis, D. A., Lunsford, E. T., & Liao, J. C. (2021). Corollary discharge enables proprioception from lateral line sensory feedback. *PLoS Biology*, *19*, e3001420. https://doi.org/10.1371/journal.pbio.3001420, PubMed: 34634044
Sperry, R. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology*, *43*, 482–489. https://doi.org/10.1037/h0055479, PubMed: 14794830
Straka, H., Simmers, J., & Chagnaud, B. P. (2018). A new perspective on predictive motor signaling. *Current Biology*, *28*, R232–R243. https://doi.org/10.1016/j.cub.2018.01.033, PubMed: 29510116
Subramanian, D., Alers, A., & Sommer, M. A. (2019). Corollary discharge for action and cognition. *Cognitive Neuroscience and Neuroimaging*, *4*, 782–790. https://doi.org/10.1016/j.bpsc.2019.05.010, PubMed: 31351985
Tian, X., Ding, N., Teng, X., Bai, F., & Poeppel, D. (2018). Imagined speech influences perceived loudness of sound. *Nature Human Behaviour*, *2*, 225–234. https://doi.org/10.1038/s41562-018-0305-8
Tian, X., & Poeppel, D. (2010). Mental imagery of speech and movement implicates the dynamics of internal forward models. *Frontiers in Psychology*, *1*, 166. https://doi.org/10.3389/fpsyg.2010.00166, PubMed: 21897822
Tian, X., & Poeppel, D. (2012). Mental imagery of speech: Linking motor and perceptual systems through internal simulation and estimation. *Frontiers in Human Neuroscience*, *6*, 314. https://doi.org/10.3389/fnhum.2012.00314, PubMed: 23226121
Tian, X., & Poeppel, D. (2013). The effect of imagination on stimulation: The functional specificity of efference copies in speech processing. *Journal of Cognitive Neuroscience*, *25*, 1020–1036. https://doi.org/10.1162/jocn_a_00381, PubMed: 23469885
Tian, X., & Poeppel, D. (2015). Dynamics of self-monitoring and error detection in speech production: Evidence from mental imagery and MEG. *Journal of Cognitive Neuroscience*, *27*, 352–364. https://doi.org/10.1162/jocn_a_00692, PubMed: 25061925
Tian, X., Zarate, J. M., & Poeppel, D. (2016). Mental imagery of speech implicates two mechanisms of perceptual reactivation. *Cortex*, *77*, 1–12. https://doi.org/10.1016/j.cortex.2016.01.002, PubMed: 26889603
Timm, J., SanMiguel, I., Saupe, K., & Schröger, E. (2013). The N1-suppression effect for self-initiated sounds is independent of attention. *BMC Neuroscience*, *14*, 2. https://doi.org/10.1186/1471-2202-14-2, PubMed: 23281832
van Boxtel, G. J. M., & Böcker, K. B. E. (2004). Cortical measures of anticipation. *Journal of Psychophysiology*, *18*, 61–76. https://doi.org/10.1027/0269-8803.18.23.61
Vercillo, T., O'Neil, S., & Jiang, F. (2018). Action-effect contingency modulates the readiness potential. *Neuroimage*, *183*, 273–279. https://doi.org/10.1016/j.neuroimage.2018.08.028, PubMed: 30114465
Vilhauer, R. P. (2017). Characteristics of inner reading voices. *Scandinavian Journal of Psychology*, *58*, 269–274. https://doi.org/10.1111/sjop.12368, PubMed: 28582587
Von Holst, E., & Mittelstaedt, H. (1950). The reafference potential. *Naturwissenschaften*, *37*, 464–476.
Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: An electric sign of sensorimotor association and expectancy in the human brain. *Nature*, *203*, 380–384. https://doi.org/10.1038/203380a0, PubMed: 14197376
Wen, W., Minohara, R., Hamasaki, S., Maeda, T., An, Q., Tamura, Y., et al. (2018). The readiness potential reflects the reliability of action consequence. *Scientific Reports*, *8*, 11865. https://doi.org/10.1038/s41598-018-30410-z, PubMed: 30089815
Whitford, T. J., Ford, J. M., Mathalon, D. H., Kubicki, M., & Shenton, M. E. (2012). Schizophrenia, myelination, and delayed corollary discharges: A hypothesis. *Schizophrenia Bulletin*, *38*, 486–494. https://doi.org/10.1093/schbul/sbq105, PubMed: 20855415
Whitford, T. J., Jack, B. N., Pearson, D., Griffiths, O., Luque, D., Harris, A. W. F., et al. (2017). Neurophysiological evidence of efference copies to inner speech. *eLife*, *6*, e28197. https://doi.org/10.7554/eLife.28197, PubMed: 29199947
Whitford, T. J., Mathalon, D. H., Shenton, M. E., Roach, B. J., Bammer, R., Adcock, R. A., et al. (2011). Electrophysiological and diffusion tensor imaging evidence of delayed corollary discharges in patients with schizophrenia. *Psychological Medicine*, *41*, 959–969. https://doi.org/10.1017/S0033291710001376, PubMed: 20663254
Whitford, T. J., Spencer, K. M., Godwin, M., Hirano, Y., Chung, L. K.-H., Vodovozov, W., et al. (2025). Gamma and theta/alpha-band oscillations in the electroencephalogram distinguish the content of inner speech. *eNeuro*, *12*, ENEURO.0297-24.2025. https://doi.org/10.1523/ENEURO.0297-24.2025, PubMed: 39843220

Wolpert, D. M., & Miall, R. C. (1996). Forward models for physiological motor control. *Neural Networks*, *9*, 1265–1279. https://doi.org/10.1016/S0893-6080(96)00035-4, PubMed: 12662535
Woods, D. L. (1995). The component structure of the N1 wave of the human auditory evoked potential. *Electroencephalography and Clinical Neurophysiology*, *44*, 102–109. PubMed: 7649012
World Medical Association. (2013). Declaration of Helsinki: Ethical principles for medical research involving human subjects. *Journal of the American Medical Association*, *310*, 2191–2194. https://doi.org/10.1001/jama.2013.281053, PubMed: 24141714
Ylinen, S., Nora, A., Leminen, A., Hakala, T., Huotilainen, M., Shtyrov, Y., et al. (2015). Two distinct auditory-motor circuits for monitoring speech production as revealed by content-specific suppression of auditory cortex. *Cerebral Cortex*, *25*, 1576–1586. https://doi.org/10.1093/cercor/bht351, PubMed: 24414279
Zivin, G. (1979). *The development of self-regulation through private speech*. New York: Wiley.