Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research Report

Real and imagined sensory feedback have comparable effects on action anticipation

Ana P. Pinheiro ^{a,b,*}, Michael Schwartze ^b, Francisco Gutiérrez-Domínguez ^a and Sonja A. Kotz ^b

^a CICPSI, Faculdade de Psicologia, Universidade de Lisboa, Lisbon, Portugal ^b Faculty of Psychology and Neuroscience, University of Maastricht, Maastricht, The Netherlands

ARTICLE INFO

Article history: Received 14 November 2019 Reviewed 21 January 2020 Revised 23 March 2020 Accepted 13 April 2020 Action editor Stephen Jackson Published online 20 May 2020

Keywords: Readiness potential Action Self-voice Imagery Forward model

ABSTRACT

The forward model monitors the success of sensory feedback to an action and links it to an efference copy originating in the motor system. The Readiness Potential (RP) of the electroencephalogram has been denoted as a neural signature of the efference copy. An open question is whether imagined sensory feedback works similarly to real sensory feedback. We investigated the RP to audible and imagined sounds in a button-press paradigm and assessed the role of sound complexity (vocal *vs.* non-vocal sound).

Sensory feedback (both audible and imagined) in response to a voluntary action modulated the RP amplitude time-locked to the button press. The RP amplitude increase was larger for actions with expected sensory feedback (audible and imagined) than those without sensory feedback, and associated with N1 suppression for audible sounds. Further, the early RP phase was increased when actions elicited an imagined vocal (self-voice) compared to non-vocal sound.

Our results support the notion that sensory feedback is anticipated before voluntary actions. This is the case for both audible and imagined sensory feedback and confirms a role of overt and covert feedback in the forward model.

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

The forward model framework explains how sensory feedback to a voluntary action is used to monitor the implementation of a motor plan (Friston, 2005; Schröger, Kotz, & SanMiguel, 2015; Wolpert & Flanagan, 2001). The forward model posits that motor regions of the brain send a copy of motor commands (the efference copy; von Holst, 1954) to sensory cortices via the cerebellum (Ito, 2008; Wolpert, Miall, & Kawato, 1998). This mechanism is inherently predictive as it instantiates expectations about feedback that are compared with actual sensory input that arises from action. In the auditory domain, concurrent expectations and sensory feedback are typically associated with attenuation of event-related potentials of the EEG (Baess, Widmann, Roye, Schröger, &



https://doi.org/10.1016/j.cortex.2020.04.030







^{0010-9452/© 2020} Elsevier Ltd. All rights reserved.

Jacobsen, 2009; Ford, Palzes, Roach, & Mathalon, 2014; Knolle, Schröger, Baess, & Kotz, 2012; Wang et al., 2014). Recent evidence indicates that sensory attenuation is modulated by changes in brain activity occurring before an action execution (Brown, Adams, Parees, Edwards, & Friston, 2013; Wang et al., 2014). Attenuation can be almost abolished when the putative generation of the efference copy is disrupted by pre-pulse transcranial magnetic stimulation (TMS) applied over the supplementary motor area (SMA) (Haggard & Whitford, 2004).

Self-initiated action is reliably preceded by voltage changes as early as one second or more before action execution. EEG signatures of neural activity preceding action may accordingly already reflect quantitative and qualitative aspects of expected sensory feedback. The Readiness Potential (RP) or Bereitschaftspotential is a slow negative deflection with a maximal amplitude over the vertex that precedes voluntary action by 1-2 s (Deecke, Heise, Kornhuber, Lang, & Lang, 1984; Di Russo et al., 2017; Shibasaki & Hallett, 2006). It comprises an early and a late phase and is related to the planning and preparation of voluntary action (Deecke et al., 1984; Di Russo et al., 2017; Reznik, Simon, & Mukamel, 2018; Shibasaki & Hallett, 2006; Vercillo, O'Neil, & Jiang, 2018). The early RP phase starts approximately 1-2 s before action onset, whereas the late RP phase begins about 500 ms before an action (Deecke et al., 1984; Di Russo et al., 2017; Reznik et al., 2018; Shibasaki & Hallett, 2006; Vercillo et al., 2018), reflecting preparatory activity in the primary motor cortex (Shibasaki & Hallett, 2006). The RP is followed by a bilateral positive potential that is observed around action onset (Deecke et al., 1984). The RP has been associated with an increase in neural activity spreading from pre-motor regions such as the SMA (early phase) to the primary motor cortex (late phase) (Ball et al., 1999; Erdler et al., 2000; Praamstra, Stegeman, Horstink, & Cools, 1996; Shibasaki & Hallett, 2006; Weilke et al., 2001; Wildgruber, Erb, Klose, & Grodd, 1997). These brain regions are connected with the cerebellum (Haggard et al., 1995), potentially supporting the central cerebellar role in the forward model (e.g., Doya, 1999; Ito, 2006; Kawato & Gomi, 1992).

Brain activity preceding an action has also been linked to specific types of action, such as grasping or touching (Gallivan, McLean, Smith, & Culham, 2011). Recent studies indicate that such activity may further represent expected feedback to voluntary action (Ford et al., 2014; Reznik et al., 2018; Vercillo et al., 2018; Wen et al., 2018). Actions with sensory effects lead to an enhanced (i.e., more negative) RP response relative to actions without sensory feedback (Reznik et al., 2018; Vercillo et al., 2018). Similarly, both early (-1000 to -500 ms) and late (-500 to 0 ms) RP phases were found to be more negative when an action was followed by a predictable rather than an unpredictable tone (Wen et al., 2018). Finally, the RP also seems to predict the neural response to a forthcoming sound: an increased RP in preparation for actions with sensory consequences was associated with a larger N1 suppression effect (e.g., Ford et al., 2014). Compared to the late RP phase, the early phase

might hence be more involved in the encoding of the expected sensory consequences of self-generated action (Reznik et al., 2018). Together, these findings suggest that increased RP responses are associated with expected sensory feedback to voluntary action.

Whether similar mechanisms are active when action is merely simulated remains an open question. For example, inner speech has been considered a special case of overt speech, which in turn constitutes our most complex motor act (Hughlings, 1958). The principle notion that perception and action can also be simulated internally is supported by ample research. Regions of the temporal cortex that respond selectively to the human voice not only activate in response to actual voice input but also when voices are imagined (Yao, Belin, & Scheepers, 2011, 2012). Recent evidence suggests that covert voice generation (potentially combining motor and sensory imagery) evokes a timelocked auditory efference copy (Ford & Mathalon, 2004; Tian & Poeppel, 2010, 2015; Ylinen et al., 2015). These findings align with evidence that inner speech activates the forward model mechanism (Stephane, Barton, & Boutros, 2001; Tian & Poeppel, 2015): the interaction between motor and sensory systems may induce speech representations without external stimulation or overt articulation (e.g., Tian & Poeppel, 2012). The forward model remains elusive as there is no clear mapping between neural signatures and forward model operations. Consequently, existing studies have inferred operations of the forward model by examining how inner speech leads to auditory N1 suppression in response to overt speech (Ford & Mathalon, 2004; Tian & Poeppel, 2015; Ylinen et al., 2015). Specifically, the production of a covert phoneme resulted in suppression of the N1 elicited by a simultaneously presented audible phoneme (Whitford et al., 2017). This effect requires the matching of content and timing of both types of phonemes (Jack et al., 2019; Whitford et al., 2017). Similar effects were observed in tactile processing: imagined selfgenerated touch was found to produce attenuation of actual tactile stimulation (Kilteni, Andersson, Houborg, & Ehrsson, 2018). These findings not only support the functional similarity of imagined and overt sensory feedback but also establish studies of imagined auditory feedback as a critical test of the forward model framework. Neural activity derived from contexts in which actual auditory feedback and proprioceptive cues are missing may help to define neural signatures for the forward model in general, and provide insight into the formation of an efference copy in particular. Following this reasoning, the current study investigates if the efference copy is generated for inner speech, and whether neural activity preceding an action may indicate how comparable anticipated imagined vs. real sensory feedback is implemented in voluntary action.

1.1. The current study

Building on previous research (Reznik et al., 2018; Vercillo et al., 2018), the primary goal of the current study was to

investigate whether real and imagined sensory feedback have comparable effects on action anticipation. We also examined whether the RP dissociates between actions eliciting vocal (self-voice) as opposed to non-vocal (computer-generated tone) sounds.

To this end, EEG was recorded from healthy participants performing a voluntary button press task either with or without sensory feedback. A well-established paradigm (Knolle et al., 2012; Knolle, Schröger, & Kotz, 2013; Pinheiro, Schwartze, & Kotz, 2018), focused on the generation of an efference copy (see also Tian & Poeppel, 2010), was used to investigate these questions. To ensure similar conditions in pre-stimulus neural activity, an overt motor action was required before both an imagined and an audible sound. Therefore, stimulus expectancy, task-relevance, and task load (motor output) were comparable across conditions. This procedure eliminates overt output in the imagery condition and provides direct evidence contrasting how expected a stimulus is in audible vs. imagined auditory feedback.

We predicted RP modulations by voluntary actions with and without sensory feedback (Fig. 1). Specifically, we expected an increased RP preparing an action that elicits a sound. This hypothesis was grounded in previous studies revealing that feedback to one's own action is estimated before sensory feedback is available (Reznik et al., 2018; Vercillo et al., 2018; Wen et al., 2018). We expected similar RP morphology preceding imagined and audible sounds, consistent with the idea that the neural mechanisms underlying the production of imagined and overt feedback, particularly speech, are the same (Feinberg, 1978; Jack et al., 2019; Tian & Poeppel, 2010, 2012; Whitford et al., 2017). The putative RP similarities would also be consistent with the notion that content-specific efference copies accompany imagined and executed actions (Kilteni et al., 2018).

We did not expect differences in the RP eliciting a selfvoice vs. a tone. Specific sensory predictions have been generated in response to simple (e.g., clicks, pure tones) or to complex and natural (e.g., speech) sounds (Baess, Jacobsen, & Schröger, 2008; Ford et al., 2014; Knolle et al., 2013, 2012; Knolle, Schwartze, Schröger, & Kotz, 2019; Martikainen, Kaneko, & Hari, 2005; Pinheiro et al., 2018). Temporally precise and content-specific predictions were reflected in N1 attenuation that did not depend on stimulus type (Knolle et al., 2019; Pinheiro et al., 2018). This suggests that the efference copy reflects not only the specific motor commands that are used to generate an action but also higher-level properties of the sensory consequences of one's own action (e.g., the phonemic properties of a self-generated sound; Niziolek, Nagarajan, & Houde, 2013). Accordingly, the RP was not expected to distinguish between different complexities of the anticipated sensory consequences of an action.

Additionally, the relationship between preparatory motor activity and sensory suppression to audible sounds was examined. We hypothesized that an increased RP would be associated with enhanced N1 suppression to self-generated audible sounds (Ford et al., 2014).

2. Method

2.1. Participants

We report how we determined our sample size,¹ all inclusion/ exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study.

Thirty-two participants took part in the EEG experiment ($M_{age} = 22.77$, SD = 4.06, age range = 18–32 years; 18 females). No data were excluded from statistical analyses. Participants were all right-handed (Oldfield, 1971), reported normal or corrected-to-normal visual acuity, normal hearing, and no history of neuropsychological dysfunction. The inclusion/ exclusion criteria were established prior to data analysis. All participants provided informed consent and were reimbursed for their time, either by course credits or a voucher.

The study was conducted in accordance with the Declaration of Helsinki and was approved by the local Ethics Committee of the University of Minho, Braga (Portugal).

2.2. Stimuli

A 680 Hz tone (50 ms duration; 70 dB sound pressure level [SPL]) and a pre-recorded self-voice speech sound (vowel/a/) were presented in separate blocks. Before the EEG experiment, a voice recording session took place: participants were instructed to repeatedly vocalize the syllable "ah". Recordings were made with an Edirol R-09 recorder and a CS-15 cardioid-type stereo microphone. After the recording, the best voice sample of the vowel/a/from each participant (i.e., constant prosody; maximum duration of 300 ms) was selected. The voice sample was edited using Audacity 2.2.2 (http://audacity. sourceforge.net) to reduce background noise and a Praat (Boersma, P., Weenink, 2013) script was applied to normalize intensity at 70 dB. The stimulus for each participant was saved in.WAV format. All voice stimuli across participants had the same duration (300 ms) and intensity (70 dB SPL).

2.3. Procedure

During the EEG recordings, participants sat comfortably at a distance of about 100 cm in front of a desktop computer monitor in a sound-attenuated and electrically shielded room. Each experimental block included three conditions (following Knolle et al., 2012, 2013): auditory-motor (AMC), auditory-only (AOC), and motor-only (MOC; Fig. 2). In the

¹ Our hypotheses were tested with mixed-effects modeling as it can lead to more efficient estimates and more powerful tests compared to the more traditional repeated-measures analysis of variance (ANOVA) (Bagiella, Sloan, & Heitjan, 2000; Jaeger, 2008). However, the field is still developing in terms of how to compute power estimates for multilevel models as several factors may impact upon this estimation, namely the number of levels, type of design, among others. Several studies have concluded that when small samples sizes are used, little to no bias is achieved in the estimates of fixed effects (e.g., Clarke & Wheaton, 2007; Maas & Hox, 2005). Notwithstanding, we note that the sample size in the current study is larger than in previous studies testing the same dependent variable (Reznik et al., 2018; Vercillo et al., 2018).



Fig. 1 – Circuitry explaining the neurofunctional mechanisms underlying action-related sensory feedback and its putative ERP correlates. Existing evidence suggests that the efference copy contains detailed information about sensory feedback, indicated by an increased RP amplitude for both audible (thin red arrows; AMC_A) and imagined (thin blue arrows; AMC_I) feedback. Actions without sensory consequences (thin dark grey arrows; MOC) are associated with decreased RP amplitude. A match between predicted and incoming actual feedback (i.e., the absence of prediction error) leads to N1 amplitude suppression. Brain regions suggested to be implicated in different operations of the forward model (efference copy formation; comparison of predicted and incoming feedback) are highlighted (grey boxes). Note: CER = cerebellum; RP = Readiness Potential; PE=Prediction Error; SMA=Supplementary Motor Area; AMC = auditory-motor condition; $AMC_A = AMC$ with audible sounds; $AMC_I = AMC$ with imagined sounds; $MOC = motor-only condition; <math>\uparrow$ increased ERP amplitude.

AMC, a button-press instantaneously elicited a tone (Experiment 1; Block 1) or an imagined tone (Experiment 1; Block 2), the prerecorded voice of the participant (Experiment 2; Block 1) or the imagined prerecorded voice of the participant (Experiment 2; Block 2; Fig. 2). Participants pressed a button approximately every 2.4 s. In the imagery blocks, participants were instructed to imagine as accurately as possible the sound (tone or self-voice) heard in the training phase (described below). Following prior studies (e.g., Tian & Poeppel, 2015), participants were trained in the imagery of their own voice and all confirmed they could induce the quasi-kinesthetic and auditory experiences vividly, without physically moving the articulators.

In the AOC, participants were instructed to attentively listen to the tones (Experiment 1) or to their pre-recorded selfvoice (Experiment 2). The acoustic stimulation from the AMC was recorded on-line and then used as the auditory sequence that was presented to the participants in the AOC. In the MOC, participants performed self-paced button presses approximately every 2.4 s but no audible or imagined tone or voice was elicited by the presses. The AMC hence necessarily always preceded the AOC (with the order of the different blocks counterbalanced across participants) but the MOC was randomized across participants.

The experimental blocks followed two training blocks. Participants performed correct taps (i.e., tapping interval no longer than 3 s and no shorter than 1.8 s – Knolle et al., 2012) in at least 75% of trials. No feedback was provided during the actual experiment. In each of the AMC and AOC blocks, 100 trials were recorded. The MOC in each experiment likewise consisted of 100 trials. The order of the experimental blocks was counterbalanced across participants. Stimulus presentation and timing was controlled by Presentation 16.3 (Neurobehavioral Systems, Inc.).² Stimuli were delivered via Sennheiser CX 300-II headphones. A BioSemi response switch (BioSemi, Amsterdam, The Netherlands) was used to record the finger taps. No part of

² 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.



Fig. 2 – Schematic illustration of the experimental design. In the auditory-motor and auditory-only Conditions, a tone or a prerecorded self-voice was presented. In the auditorymotor Condition, actions could lead to audible (AMC_A) vs. imagined (AMC_I) sensory consequences. Note: AMC = auditory-motor condition; AMC_A = AMC with audible sounds; AMC_I = AMC with imagined sounds.

the study procedures was pre-registered prior to the research being conducted.

2.4. EEG data acquisition and analysis

EEG data were recorded using a 64-channel BioSemi Active Two system (BioSemi, Amsterdam, The Netherlands) in a continuous mode at a digitization rate of 512 Hz and stored on hard disk for later analysis.

Data were re-referenced to the average of the left and right mastoids using Brain Vision Analyzer 2 software (Brain Products, Gilching, Germany). Data were then filtered with a .1-100 Hz bandpass filter (zero phase shift Butterworth, order 2) and then segmented into epochs that were time-locked to the motor response, from -1500 ms to 500 ms, with a 500 ms baseline (-1500 ms to -1000 ms). Tapping intervals shorter than 1.8 sec were considered errors and the respective trials were excluded from all further analyses. Independent component analyses (Infomax ICA) were performed to remove stereotyped noise (mainly ocular movements and blinks) by subtracting the corresponding components (Jung et al., 1998; Mennes, Wouters, Vanrumste, Lagae, & Stiers, 2010). After segmentation, epochs with amplitudes exceeding ±100 µV were automatically rejected from further analysis. Individual ERPs were averaged separately for each condition. Grand averages waveforms were calculated for each condition (AMC, MOC), task (eliciting

an audible vs. imagined sound), and stimulus type (tone vs. self-voice).

Following previous studies (e.g., Vercillo et al., 2018; Wen et al., 2018), RP amplitude analyses were performed at Cz and locked to the time of the button press. An early RP phase was examined from -1000 to -500 ms before button press onset, whereas a late RP phase was examined from -500 ms until button press onset (Reznik et al., 2018; Vercillo et al., 2018; Wen et al., 2018). ERP amplitude was quantified as the mean voltage within these segments. Trials included in the analyses converged in terms of their physical features (all trials involved pressing a button that could either be followed or not followed by a sound). The only difference between the conditions was the action consequence.

The analysis of auditory potentials (time-locked to sound onset), in response to self-generated audible sounds, was controlled for motor activity by computing difference waveforms between AMC and MOC. The mean N1 amplitude for self-generated (AMC) and externally generated (AOC) audible sounds was calculated in a time window of 70–110 ms post-sound onset following prior studies (e.g., Pinheiro et al., 2018).

2.5. Statistical data analyses

ERP amplitudes were analyzed by means of mixed-linear models using the lmer4 (Bates, Maechler, Bolker, & Walker, 2015) and ImerTest (Kuznetsova, Brockhoff, & Christensen, 2016) packages in the R environment (R3.4.3. GUI 1.70) used to estimate fixed and random coefficients. In contrast to traditional repeated-measures ANOVA analyses. LMER allows controlling for the variance associated with random factors such as random effects for participants in terms of ERP magnitude (Baayen, Davidson, & Bates, 2008). Amplitude was included as outcome, whereas participants and hallucination proneness³ (total score of the Launay Slade Hallucination Scale - Castiajo & Pinheiro, 2017) were included as random effects. The role of the fixed effects of condition (AMC, MOC), task (audible feedback [AMC_A], imagined feedback [AMC_I]), and stimulus type (self-voice, tone) was probed in two different models.⁴

Moreover, we tested for differences in the pattern of descend of the electric potential (from the early to the late phase) across conditions using mixed linear models, including the same fixed effects.

3. Results

3.1. Behavioral results

At the behavioral level, there were no significant differences between button presses that were either associated or not

³ Hallucination proneness (measured with the Launay-Slade Hallucination Scale – Castiajo & Pinheiro, 2017) was included as random effect considering that it may affect sensory prediction, namely the N1 suppression to self-generated sounds (Pinheiro et al., 2018).

⁴ Analysis code is presented as Supplementary Material.

associated with an audible or an imagined sound. The average tapping interval was 2744.61 ms (SD = 475.06 ms) in the MOC, 2668.97 ms (SD = 324.22 ms) in the AMC_A condition for tones, 2686.99 ms (SD = 395.73 ms) in the AMC_I condition for tones, 2552.92 ms (SD = 270.43 ms) in the AMC_A condition for self-voices, and 2771.60 ms (SD = 485.11 ms) in the AMC_I condition for self-voices. In all cases, the Kolmogorov-Smirnov test revealed that the data fit a normal distribution (p > .05). The overall accuracy (i.e., tapping interval no longer than 3 s and no shorter than 1.8 s) was above 98% (MOC: 99.19%; AMC_A - Tone: 99.16%; AMC_I - Tone: 98.81%; AMC_A - Self-voice: 99.47%; AMC_I - Self-voice: 99.75%). Results did not show accuracy differences between conditions (p > .05).

3.2. ERP results

Fig. 3 illustrates the RP and its topographical distribution preceding actions with *vs.* without sensory effects.

3.2.1. Readiness Potential

A slope analysis revealed that the RP was characterized by a more negative deflection in the late compared to the early phase (β = -.667, SE = .282, t (134) = -2.367, p = .019, 95% CI [-1.219, -.115]). The pattern of RP descend did not differ significantly between conditions (p > .05 for all comparisons; Fig. 4).

The first model examined whether actions leading to sensory consequences or not were associated with RP differences, by testing the effects of condition (AMC_A, AMC_I, MOC) irrespective of stimulus type. Consistent with our hypothesis, brain activity preceding button presses that elicited audible ($\beta = -.278$, SE = .065, t (134) = -4.250, p < .001, 95% CI [-.406, -.150]) or imagined ($\beta = -.233$, SE = .068, t (134) = -3.448, p < .001, 95% CI [-.366, -.101]) sounds was characterized by a similarly increased (i.e., more negative) early RP phase compared to button presses without sensory effects (Figs. 3 and 5). There were no differences in the early RP preceding overt or imagined sounds (p = .430).

In the late phase, the RP enhancement before actions with sensory outcomes (AMC vs. MOC) was stronger when the button press elicited an imagined sound compared to an audible sound (β = -.508, SE = .234, t (127) = -2.175, *p* = .031, 95% CI [-.966, -.050]; Figs. 3 and 5). There were no differences in the late RP preceding overt or imagined sounds (*p* = .373).

The second model examined whether actions leading to sensory consequences (AMC_A, AMC_I) were associated with RP differences as a function of stimulus type. We observed that the early RP phase was less negative before actions that elicited an imagined tone than before actions that elicited an imagined voice ($\beta = .756$, SE = .271, t (96) = 2.796, p = .006, 95% CI [.226, 1.286]; Fig. 5). The RP did not differ before actions that elicited an overt tone compared to an overt voice (p = .750).



Fig. 3 – Grand average waveforms (at electrode Cz) showing the early (-1000 to -500 ms; dark grey) and late (-500 ms relative to action onset; light grey) RP phases preceding voluntary actions with (AMC_A and AMC_I) vs. without (MOC) sensory effects. Topographical maps illustrate the scalp distribution of the RP for each condition (AMC vs. MOC), task (audible vs. imagined sound), and stimulus type (self-voice; tone) and are depicted for the early (-1000 to -500 ms time window) and the late phases (successive 100 ms time windows from -500 ms until button press onset). Note: AMC = auditory-motor condition; AMC_A = AMC with audible sounds; AMC_I = AMC with imagined sounds; MOC = motor-only condition. To further illustrate temporal differences in the topographical distribution of the RP across conditions, topo maps are shown for successive 100 ms time windows in the late RP phase.



Fig. 4 – ERP slopes for the early phase of the RP (-1000 to -500 ms) as a function of condition (AMC, MOC), task (audible feedback, imagined feedback), and stimulus type (tone, voice). Note: AMC = auditory-motor condition; AMC_A = AMC with audible sounds; AMC_I = AMC with imagined sounds; MOC = motor-only condition. The shaded area represents 95% confidence intervals.

3.2.2. The relationship between the RP and the auditory N1 We performed an additional analysis to explore the relation of the RP and the N1 response to self-generated audible sounds. Generation of an efference copy has been assumed to reduce the sensitivity of the auditory cortex to unaltered auditory feedback (e.g., Behroozmand & Larson, 2011; Ventura, Nagarajan, & Houde, 2009). Accordingly, an increased RP before actions leading to sensory consequences should be mirrored by enhanced N1 suppression to audible



Fig. 5 – Mean ERP activity in the early and late phases of the RP as a function of condition (AMC, MOC), task (audible feedback, imagined feedback), and stimulus type (tone, voice). Note: $AMC_A = AMC$ with audible sounds; $AMC_I = AMC$ with imagined sounds; MOC = motor-only condition; RP = Readiness Potential. Error bars represent standard error.

sounds, i.e., less negative N1 to self-generated (AMC) compared to externally generated (AOC) sounds. Consistently, N1 suppression was predicted by the late RP response, irrespective of stimulus type (i.e., self-voice vs. tone): the larger the RP in preparation for actions eliciting a sound, the larger the N1 suppression (β = .368, SE = .161, t = 2.283, p = .027, 95% CI [.052, .684]; Fig. 6).

4. Discussion

The forward model framework postulates a functional relation of actions and their sensory consequences. According to this framework, brain activity preceding an action such as pressing a button that elicits a sound reflects the generation of an inherently predictive efference copy (Reznik et al., 2018; Roussel, Hughes, & Waszak, 2013; Vercillo et al., 2018), consistent with a role of the motor system in sensory prediction (Wolpert & Flanagan, 2001). The current study set out to replicate the finding that preparatory motor activity is modulated by anticipating the sensory effects elicited by an action: button presses eliciting imagined or audible sounds were associated with a stronger negativity compared to button presses with no auditory consequences. We then tested and confirmed the hypothesis that motor preparation for imagined sounds is comparable to motor preparation for overt sounds: both were associated with enhanced early and late phases of the RP. Furthermore, the early RP was sensitive to stimulus type (self-voice vs. tone) in the imagery task. This supports the view that a representation of the expected sensory feedback is encoded in the RP, confirming the role of the RP as electrophysiological marker of action-related predictive mechanisms (Reznik et al., 2018; Vercillo et al., 2018).

4.1. Preparing for actions that elicit audible vs. imagined sounds

The observed differences in the RP preceding actions with vs. without sensory outcomes (i.e., increased RP in the AMC compared to MOC) are in line with recent reports concerning modulatory effects of the anticipation of sensory feedback on brain activity prior to action execution (Reznik et al., 2018; Vercillo et al., 2018). More specifically, the RP does not reflect motor preparation per se but is modulated by the action meaning and sensory consequences (Di Russo et al., 2017). Of note, accuracy data showed no differences between conditions (AMC, MOC), tasks (audible vs. imagined sound), and stimulus types (self-voice vs. tone). This suggests that the type of sensory outcome of an action does not affect action execution: only brain activity preceding an action was differentially modulated by type of feedback. These findings are consistent with a system that estimates the sensory consequences of an action before feedback is available (Reznik et al., 2018; Vercillo et al., 2018; Wen et al., 2018), with direct implications for sensory responsiveness to self-generated feedback. Sensory attenuation depends on the degree of convergence between expected and incoming sensory feedback: a larger RP was associated with a stronger N1 attenuation to feedback (Reznik et al., 2018; Vercillo et al., 2018), as



Fig. 6 – Grand average waveforms contrasting self-triggered and externally triggered self-voice (Panel A) and tone (Panel B) at Cz. Relationship between the RP and the N1 suppression to audible sounds at Cz (Panel C). *Note:* AMC = auditory-motor condition; AOC = auditory-only condition. In Panel C, more negative values indicate less negative amplitude in response to self-initiated (AMC) sounds compared to externally triggered (AOC) sounds. The shaded area represents 95% confidence intervals.

confirmed by the current study. The SMA has been pointed out as a likely source of the modifying activity on sensory cortices (Reznik, Ossmy, & Mukamel, 2015). Accordingly, TMS over the SMA almost abolished sensory suppression effects following voluntary actions (Haggard & Whitford, 2004).

Confirming our predictions, the current study revealed an increased (i.e., more negative) RP preparing for button presses that elicited either an audible or an imagined sound relative to actions with no sensory effects. When compared to audible speech, inner speech was also found to be associated with a precise and content-specific efference copy that contributed to sensory attenuation in the absence of any actual action (Jack et al., 2019; Whitford et al., 2017). Other studies showed that similar brain regions are activated when contrasting inner and overt speech, including the SMA (Tian, Zarate, & Poeppel, 2016), which is also known to be a main source to the RP (Ball et al., 1999; Erdler et al., 2000; Praamstra et al., 1996; Shibasaki & Hallett, 2006; Weilke et al., 2001; Wildgruber et al., 1997). This suggests that the two processes are functionally equivalent.

The RP enhancement for actions with sensory outcomes (AMC vs. MOC) was stronger when the button press triggered an imagined compared to an audible sound in the late RP phase. However, this enhancement did not depend on the task (audible vs. imagined sound) in the early RP phase. Compared to the later RP phase, the early phase seems to be involved to a greater extent in the encoding of predicted

auditory consequences of self-generated actions (Reznik et al., 2018) and engages higher-order motor areas, such as the SMA (Shibasaki & Hallett, 2006). A sensorimotor function has been assigned to the SMA, which makes use of motor programs to anticipate sensory consequences that may optimize perception (Lima, Krishnan, & Scott, 2016). Previous studies also found a link between the early RP phase and behavioral reports of intentional binding (Jo, Wittmann, Hinterberger, & Schmidt, 2014), with implications for a sense of agency (Kranick et al., 2013; Kühn, Brass, & Haggard, 2013; Moore, Ruge, Wenke, Rothwell, & Haggard, 2010). The current findings suggest similar mechanisms for the anticipation of audible vs. imagined action outcomes. They are in line with studies showing that mental imagery in general, and inner speech in particular also engage the forward model (Grush, 2004; Sirigu et al., 2006; Tian & Poeppel, 2012). Similar neural populations were proposed to underlie the representation of auditory efference copies for overt and covert speech (Tian & Poeppel, 2013), suggesting that inner and over speech rely on similar underlying neural processes (Alderson-Day & Fernyhough, 2015). In the specific context of imagery, estimates of the sensory consequences of actions may underlie the subjective experience of 'hearing' (Lima et al., 2016).

Paradigms that include mental imagery have the advantage of eliminating overt output, proprioceptive cues, and their artifacts in the EEG signal. The current findings confirm that mental imagery can be a valid method in research probing action-perception interactions in the forward model framework (e.g., Tian & Poeppel, 2015).

4.2. Preparing for actions that elicit a vocalization vs. a tone

When testing the effects of stimulus type on the RP, our hypothesis was not confirmed. We observed that the early RP phase was modulated by stimulus type (self-voice vs. tone) but only when actions were followed by sound imagery: motor preparation for actions that triggered an imagined voice were associated with an increased negativity compared to actions that triggered an imagined tone.

Specific sensory predictions (both temporally precise and content-specific) have been generated irrespective of stimulus complexity and showed in similar N1 suppression for simple (e.g., sinusoidal sounds) and complex (e.g., speech sounds) audible sounds (e.g., Knolle et al., 2019; Pinheiro et al., 2018). The N1 suppression effects suggest that the efference copy for a self-generated speech sound generates a precise prediction concerning the temporal and acoustic appearance of the feedback (Jack et al., 2019; Knolle et al., 2019; Whitford et al., 2017).

However, we acknowledge potential differences between the imagery of one's own voice vs. a sinusoidal tone. Even though imagining to hear one's own voice and imagining to hear a sinusoidal tone both engage memory-based mechanisms (Tian & Poeppel, 2012), self-voice imagery might be more strongly mediated by motor-based mechanisms (Ma & Tian, 2019; Tian, Ding, Teng, Bai, & Poeppel, 2018). Previous studies revealed that inner speech is affected by motor articulation, i.e. it is not independent of the movements that a person would use in overt speech (e.g., Oppenheim & Dell, 2010). Furthermore, inner speech activated brain regions involved in motor planning (Barch et al., 1999; Yetkin et al., 1995), suggesting motor simulation. Stronger motor-to-auditory links are therefore expected when imagining a self-vocalization than when imagining a pure tone. Accordingly, imagining one's one voice could represent an example of combined motor (articulation) and sensory (hearing) imagery, whereas imagining a tone could represent an example of primarily sensory imagery. It is plausible that differences in the early RP phase before actions that elicited an imagined voice vs. a tone are related to differences in the predictions generated by motor- (self-voice imagery) and memory-based (tone imagery) mechanisms. This possibility needs to be tested in future studies.

4.3. Implications

Together, the current findings confirm the contribution of motor areas (the neural generators of the RP) to the encoding of the expectation of sensory consequences of one's own actions, potentially by modulating activity in sensory cortices (Lima et al., 2016). This mechanism could explain enhanced N1 suppression to self-generated (audible) feedback. Future studies should more clearly separate motor and sensory imagery processes by providing imagery instructions that more directly emphasize the generative aspect of inner speech (i.e., articulation imagery of own voice/other voice vs. hearing imagery of own voice/other voice) to understand how differences in motor involvement during imagery affect the RP. In addition, it remains to be specified how the RP is affected by different degrees of predictability regarding an upcoming sound (e.g., trials in which the received feedback violates the prediction).

Sensory predictions generated in preparation for voluntary actions are thought to contribute to the discrimination of selfgenerated and externally triggered sensory stimuli (Ford & Mathalon, 2004; Friston, 2012). Alterations in these mechanisms may contribute to altered self-monitoring and feedback control, and specifically lead to misperceptions of the origin of self-generated thoughts and actions in auditory verbal hallucinations (e.g., Pinheiro et al., 2018). In agreement with this hypothesis, N1 suppression effects are reduced or even absent for speaking and inner speech in schizophrenia patients experiencing AVH (Ford et al., 2001; Tian & Poeppel, 2012) and also in nonclinical voice hearers (Pinheiro et al., 2018), who tend to erroneously identify their own voice as generated by somebody else (Allen et al., 2004; Pinheiro, Farinha-Fernandes, Roberto, & Kotz, 2019; Pinheiro, Rezaii, Rauber, & Niznikiewicz, 2016).

5. Conclusions

A robust body of evidence suggests that motor regions of the brain are involved in the encoding of expectations regarding action-related sensory feedback. By showing an increased RP before actions leading to sensory consequences (vs. actions without sensory effects), the current study confirms that sensory predictions are generated prior to voluntary action (Vercillo et al., 2018) and that these predictions modulate the sensory responsiveness to sound. By showing similarities in preparatory motor activity before actions that lead to audible or to imagined sounds, the current study furthermore supports the notion of a functional equivalence between auditory imagery and overt auditory perception. In contrast, differential modulation of the RP preceding actions that trigger an imagined self-voice vs. a tone suggest differences in the precision of the efference copy that depend on the nature of the imagined sound (combined motor and sensory imagery in the case of the self-voice vs. primarily sensory imagery in the case of the tone). The experimental design that has been applied in the current study thus offers a promising new approach to examine the forward model on the basis of imagined sounds in general, and inner speech in particular.

Credit author statement

APP, MS, and SK designed the study and wrote the protocol. APP, MS, and SK managed the literature searches. APP and FG collected and analyzed the data. APP and FG undertook the

statistical analyses. APP wrote the first draft of the manuscript. All authors contributed to and have approved the final manuscript.

Acknowledgments

This work was supported by the Portuguese Science and Technology Foundation (Fundação para a Ciência e a Tecnologia [FCT]; grant numbers IF/00334/2012, PTDC/MHC-PCN/ 0101/2014). The Authors gratefully acknowledge all the participants who collaborated in the study, and particularly Dr. Franziska Knolle for feedback on stimulus generation and Dr. Magda Sofia Roberto for her advice on mixed linear models.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cortex.2020.04.030.

REFERENCES

- Alderson-Day, B., & Fernyhough, C. (2015). Inner speech: Development, cognitive functions, phenomenology, and neurobiology. Psychological Bulletin, 141, 931–965.
- Allen, P., Johns, L. C., Fu, C. H. Y., Broome, M. R., Vythelingum, G. N., & McGuire, P. K. (2004). Misattribution of external speech in patients with hallucinations and delusions. Schizophrenia Research, 69(2–3), 277–287. https://doi.org/ 10.1016/j.schres.2003.09.008.
- 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.
- Baess, P., Jacobsen, T., & Schröger, E. (2008). Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: Evidence for internal forward models with dynamic stimulation. International Journal of Psychophysiology, 70(2), 137–143. https://doi.org/10.1016/ j.ijpsycho.2008.06.005.
- Baess, P., Widmann, A., Roye, A., Schröger, E., & Jacobsen, T. (2009). Attenuated human auditory middle latency response and evoked 40-Hz response to self-initiated sounds. European Journal of Neuroscience, 29(7), 1514–1521. https://doi.org/ 10.1111/j.1460-9568.2009.06683.x.
- Bagiella, E., Sloan, R. P., & Heitjan, D. F. (2000). Mixed-effects models in psychophysiology. Psychophysiology, 37, 13–20. https://doi.org/10.1017/S0048577200980648.
- Ball, T., Schreiber, A., Feige, B., Wagner, M., Lücking, C. H., & Kristeva-Feige, R. (1999). The role of higher-order motor areas in voluntary movement as revealed by high-resolution EEG and fMRI. Neuroimage, 10(6), 682–694. https://doi.org/10.1006/ nimg.1999.0507.
- Barch, D. M., Sabb, F. W., Carter, C. S., Braver, T. S., Noll, D. C., & Cohen, J. D. (1999). Overt verbal responding during fMRI scanning: Empirical investigations of problems and potential solutions. Neuroimage, 10, 642–657. https://doi.org/10.1006/ nimg.1999.0500.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Ime4: linear mixed-effects models using Eigen and S4. Journal of Statistical Software, 67, 1–48. https://doi.org/10.18637/ jss.v067.i01.

- Behroozmand, R., & Larson, C. R. (2011). Error-dependent modulation of speech-induced auditory suppression for pitchshifted voice feedback. BMC Neuroscience, 12(1), 54. https:// doi.org/10.1186/1471-2202-12-54.
- Boersma, P., & Weenink, D. (2013). Praat: Doing phonetics by computer [Computer program]. Version 5.3.53. Retrieved from http://www.praat.org/.
- Brown, H., Adams, R. A., Parees, I., Edwards, M., & Friston, K. (2013). Active inference, sensory attenuation and illusions. *Cognitive Processing*, 14(4), 411–427. https://doi.org/10.1007/ s10339-013-0571-3.
- Castiajo, P., & Pinheiro, A. P. (2017). On "hearing" voices and "seeing" things: Probing hallucination predisposition in a Portuguese nonclinical sample with the Launay-Slade Hallucination Scale-revised. Frontiers in Psychology, 8, 1138. https://doi.org/10.3389/fpsyg.2017.01138.
- Clarke, P., & Wheaton, B. (2007). Addressing data sparseness in contextual population research: Using cluster analysis to create synthetic neighborhoods. Sociological Methods and Research, 35(3), 311–351. https://doi.org/10.1177/ 0049124106292362.
- Deecke, L., Heise, B., Kornhuber, H. H., Lang, M., & Lang, W. (1984). Brain potentials associated with voluntary manual tracking: Bereitschaftspotential, conditioned premotion positivity, directed attention potential, and relaxation potential: Anticipatory activity of the limbic and frontal cortex. Annals of the New York Academy of Sciences, 425, 450–464. https://doi.org/ 10.1111/j.1749-6632.1984.tb23567.x.
- Di Russo, F., Berchicci, M., Bozzacchi, C., Perri, R. L., Pitzalis, S., & Spinelli, D. (2017). Beyond the "Bereitschaftspotential": Action preparation behind cognitive functions. *Neuroscience and Biobehavioral Reviews*, 78, 57–81. https://doi.org/10.1016/ j.neubiorev.2017.04.019.
- Doya, K. (1999). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Networks*, 12(7–8), 961–974. https://doi.org/10.1016/S0893-6080(99)00046-5.
- Erdler, M., Beisteiner, R., Mayer, D., Kaindl, T., Edward, V., Windischberger, C., et al. (2000). Supplementary motor area activation preceding voluntary movement is detectable with a whole-scalp magnetoencephalography system. *Neuroimage*, 11(6 Pt 1), 697–707. https://doi.org/10.1006/ nimg.2000.0579.
- Feinberg, I. (1978). Efference copy and corollary discharge: Implications for thinking and its disorders*. Schizophrenia Bulletin, 4(4), 636–640. https://doi.org/10.1093/schbul/4.4.636.
- Ford, J. M., & Mathalon, D. H. (2004). Electrophysiological evidence of corollary discharge dysfunction in schizophrenia during talking and thinking. *Journal of Psychiatric Research*, 38(1), 37–46. https://doi.org/10.1016/S0022-3956(03)00095-5.
- Ford, J. M., Mathalon, D. H., Kalba, S., Whitfield, S., Faustman, W. O., & Roth, W. T. (2001). Cortical responsiveness during inner speech in schizophrenia: An event-related potential study. American Journal of Psychiatry, 158(11), 1914–1916. https://doi.org/10.1176/appi.ajp.158.11.1914.
- Ford, J. M., Palzes, V. A., Roach, B. J., & Mathalon, D. H. (2014). Did i do that? Abnormal predictive processes in schizophrenia when button pressing to deliver a tone. Schizophrenia Bulletin, 40(4), 804–812. https://doi.org/10.1093/schbul/sbt072.
- Friston, K. (2005). A theory of cortical responses. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 360(1456), 815–836. https://doi.org/10.1098/ rstb.2005.1622.
- Friston, K. (2012). Prediction, perception and agency. International Journal of Psychophysiology, 83(2), 248–252. https://doi.org/ 10.1016/j.ijpsycho.2011.11.014.
- Gallivan, J. P., McLean, D. A., Smith, F. W., & Culham, J. C. (2011). Decoding effector-dependent and effector-independent movement intentions from human parieto-frontal brain

activity. Journal of Neuroscience, 31(47), 17149–17168. https://doi.org/10.3390/s150613406.

- Grush, R. (2004). The emulation theory of representation: Motor control, imagery, and perception. Behavioral and Brain Sciences, 27(3), 377–396. https://doi.org/10.1017/s0140525x04000093.
- Haggard, P., Miall, R. C., Wade, D., Fowler, S., Richardson, A., Anslow, P., et al. (1995). Damage to cerebellocortical pathways after closed head injury: A behavioural and magnetic resonance imaging study. *Journal of Neurology, Neurosurgery,* and Psychiatry, 58, 433–438.
- Haggard, P., & Whitford, B. (2004). Supplementary motor area provides an efferent signal for sensory suppression. Cognitive Brain Research, 19(1), 52–58. https://doi.org/10.1016/ j.cogbrainres.2003.10.018.
- von Holst, E. (1954). Relations between the central Nervous System and the peripheral organs. The British Journal of Animal Behaviour, 2(3), 89–94. https://doi.org/10.1016/S0950-5601(54) 80044-X.
- Hughlings, J. J. (1958). Selected writings of John Hughlings Jackson. New York, NY: Basic Books.
- Ito, M. (2006). Cerebellar circuitry as a neuronal machine. Progress in Neurobiology, 78(3–5), 272–303. https://doi.org/10.1016/ j.pneurobio.2006.02.006.
- Ito, M. (2008). Control of mental activities by internal models in the cerebellum. Nature Reviews Neuroscience, 9, 304–313. https://doi.org/10.1038/nrn2332.
- 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.
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. Journal of Memory and Language, 59(4), 434–446. https://doi.org/ 10.1016/j.jml.2007.11.007.
- 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.
- Jung, T.-P., Humphries, C., Lee, T.-W., Makeig, S., McKeown, M. J., Iragui, V., et al. (1998). Extended ICA removes artifacts from electroencephalographic recordings. Advances in Neural Information Processing Systems, 10, 894–900.
- Kawato, M., & Gomi, H. (1992). A computational model of four regions of the cerebellum based on feedback-error learning. Biological Cybernetics, 68, 95–103. https://doi.org/10.1007/ BF00201431.
- Kilteni, K., Andersson, B. J., Houborg, C., & Ehrsson, H. H. (2018). Motor imagery involves predicting the sensory consequences of the imagined movement. *Nature Communications*, 9, 1617. https://doi.org/10.1038/s41467-018-03989-0.
- 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 selfinitiated vowels. Frontiers in Neuroscience, 13, 1146. https:// doi.org/10.3389/fnins.2019.01146.
- Kranick, S. M., Moore, J. W., Yusuf, N., Martinez, V. T., Lafaver, K., Edwards, M. J., et al. (2013). Action-effect binding is decreased in motor conversion disorder: Implications for sense of agency. Movement Disorders, 28(8), 1110–1116. https://doi.org/ 10.1002/mds.25408.

- Kühn, S., Brass, M., & Haggard, P. (2013). Feeling in control: Neural correlates of experience of agency. Cortex, 49(7), 1935–1942. https://doi.org/10.1016/j.cortex.2012.09.002.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). ImerTest: Tests in linear mixed effects models. R Package Version 2.0-33. Retrieved from https://cran.r-project.org/ package=lmerTest.
- Lima, C. F., Krishnan, S., & Scott, S. K. (2016). Roles of supplementary motor areas in auditory processing and auditory imagery. *Trends in Neurosciences*, 39(8), 527–542. https://doi.org/10.1016/j.tins.2016.06.003.
- Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. Methodology, 1(3), 86–92. https://doi.org/ 10.1027/1614-2241.1.3.86.
- Martikainen, M. H., Kaneko, K. I., & Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. *Cerebral Cortex*, 15(3), 299–302. https://doi.org/10.1093/ cercor/bhh131.
- Ma, O., & Tian, X. (2019). Distinct mechanisms of imagery differentially influence speech perception. ENeuro, 6(5). https://doi.org/10.1523/ENEURO.0261-19.2019. ENEURO.0261-19.2019.
- Mennes, M., Wouters, H., Vanrumste, B., Lagae, L., & Stiers, P. (2010). Validation of ICA as a tool to remove eye movement artifacts from EEG/ERP. Psychophysiology, 47(6), 1142–1150. https://doi.org/10.1111/j.1469-8986.2010.01015.x.
- Moore, J. W., Ruge, D., Wenke, D., Rothwell, J., & Haggard, P. (2010). Disrupting the experience of control in the human brain: Pre-supplementary motor area contributes to the sense of agency. Proceedings of the Royal Society B: Biological Sciences, 277(1693), 2503–2509. https://doi.org/10.1098/ rspb.2010.0404.
- 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.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4.
- Oppenheim, G. M., & Dell, G. S. (2010). Motor movement matters: The flexible abstractness of inner speech. *Memory & Cognition*, 38(8), 1147–1160. https://doi.org/10.3758/MC.38.8.1147.
- Pinheiro, A. P., Farinha-Fernandes, A., Roberto, M. S., & Kotz, S. (2019). Self-voice perception and its relationship with hallucination predisposition. *Cognitive Neuropsychiatry*, 24(4), 237–255.
- Pinheiro, A. P., Rezaii, N., Rauber, A., & Niznikiewicz, M. (2016). Is this my voice or yours? The role of emotion and acoustic quality in self-other voice discrimination in schizophrenia. *Cognitive Neuropsychiatry*, 21(4), 335–353. https://doi.org/ 10.1080/13546805.2016.1208611.
- Pinheiro, A. P., Schwartze, M., & Kotz, S. A. (2018). Voice-selective prediction alterations in nonclinical voice hearers. *Scientific Reports*, 8(1), 14717. https://doi.org/10.1038/s41598-018-32614-9.
- Praamstra, P., Stegeman, D. F., Horstink, M. W. I. M., & Cools, A. R. (1996). Dipole source analysis suggests selective modulation of the supplementary motor area contribution to the readiness potential. Electroencephalography and Clinical Neurophysiology, 98, 468–477.
- Reznik, D., Ossmy, O., & Mukamel, R. (2015). Enhanced auditory evoked activity to self-generated sounds is mediated by primary and supplementary motor cortices. *Journal of Neuroscience*, 35(5), 2173–2180. https://doi.org/10.1523/ JNEUROSCI.3723-14.2015.
- 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.

Roussel, C., Hughes, G., & Waszak, F. (2013). A preactivation account of sensory attenuation. *Neuropsychologia*, 51(5), 922–929. https://doi.org/10.1016/ j.neuropsychologia.2013.02.005.

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.

Shibasaki, H., & Hallett, M. (2006). What is the Bereitschaftspotential? Clinical Neurophysiology, 117(11), 2341–2356. https://doi.org/10.1016/j.clinph.2006.04.025.

Sirigu, A., Duhamel, J.-R., Cohen, L., Pillon, B., Dubois, B., & Agid, Y. (2006). The mental representation of hand movements after parietal cortex damage. *Science*, 273(5281), 1564–1568. https://doi.org/10.1126/science.273.5281.1564.

Stephane, M., Barton, S., & Boutros, N. N. (2001). Auditory verbal hallucinations and dysfunction of the neural substrates of speech. Schizophrenia Research, 50(1–2), 61–78. https://doi.org/ 10.1016/S0920-9964(00)00150-X.

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.

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.

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(7), 1020–1036. https://doi.org/10.1162/jocn_a_00381.

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(2), 352–364. https://doi.org/10.1162/jocn_a_00692.

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.

Ventura, M. I., Nagarajan, S. S., & Houde, J. F. (2009). Speech target modulates speaking induced suppression in auditory cortex. BMC Neuroscience, 10, 58. https://doi.org/10.1186/1471-2202-10-58.

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.

Wang, J., Mathalon, D. H., Roach, B. J., Reilly, J., Keedy, S. K., Sweeney, J. A., et al. (2014). Action planning and predictive coding when speaking. Neuroimage, 91, 91–98. https://doi.org/ 10.1016/j.neuroimage.2014.01.003.

Weilke, F., Spiegel, S., Boecker, H., von Einsiedel, H. G., Conrad, B., Schwaiger, M., et al. (2001). Time-resolved fMRI of activation patterns in M1 and SMA during complex voluntary movement. *Journal of Neurophysiology*, 85(5), 1858–1863. https://doi.org/ 10.1152/jn.2001.85.5.1858.

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.

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.

Wildgruber, D., Erb, M., Klose, U., & Grodd, W. (1997). Sequential activation of supplementary motor area and primary motor cortex during self-paced finger movement in human evaluated by functional MRI. *Neuroscience Letters*, 227(3), 161–164. https://doi.org/10.1016/S0304-3940(97)00329-7.

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. Current Biology, 11(18), R729–R732. https://doi.org/10.1016/S0960-9822(01)00432-8.

Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. Trends in Cognitive Sciences, 2, 338–347. https://doi.org/10.1016/S1364-6613(98)01221-2.

Yao, B., Belin, P., & Scheepers, C. (2011). Silent reading of direct versus indirect speech activates voice-selective areas in the auditory cortex. *Journal of Cognitive Neuroscience*, 23(10), 3146–3152. https://doi.org/10.1162/jocn_a_00022.

Yao, B., Belin, P., & Scheepers, C. (2012). Brain "talks over" boring quotes: Top-down activation of voice-selective areas while listening to monotonous direct speech quotations. *Neuroimage*, 60(3), 1832–1842. https://doi.org/10.1016/ j.neuroimage.2012.01.111.

Yetkin, F. Z., Hammeke, T. A., Swanson, S. J., Morris, G. L., Mueller, W. M., McAuliffe, T. L., et al. (1995). A comparison of functional MR activation patterns during silent and audible language tasks. American Journal of Neuroradiology, 16, 1087–1092.

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 contentspecific suppression of auditory cortex. *Cerebral Cortex*, 25(6), 1576–1586. https://doi.org/10.1093/cercor/bht351.