



Registered Report

Distinct neural mechanisms support inner speaking and inner hearing[☆]



Ladislav Nalborczyk^{a,b,*}, Marieke Longcamp^b, Mireille Bonnard^c,
Victor Serveau^a, Laure Spieser^b and F.-Xavier Alario^a

^a Aix Marseille Univ, CNRS, LPC, Marseille, France

^b Aix Marseille Univ, CNRS, LNC, Marseille, France

^c Aix Marseille Univ, INSERM, INS, Marseille, France

ARTICLE INFO

Article history:

Protocol received: 22 November 2021

Protocol approved: 13 April 2022

Received 24 May 2023

Reviewed 21 June 2023

Revised 11 September 2023

Accepted 13 September 2023

Action editor Robert D. McIntosh

Published online 10 October 2023

Keywords:

Inner speech

Inner speaking

Inner hearing

Transcranial magnetic stimulation

Motor evoked potential

Cortical excitability

ABSTRACT

Humans have the ability to mentally examine speech. This covert form of speech production is often accompanied by sensory (e.g., auditory) percepts. However, the cognitive and neural mechanisms that generate these percepts are still debated. According to a prominent proposal, inner speech has at least two distinct phenomenological components: inner speaking and inner hearing. We used transcranial magnetic stimulation to test whether these two phenomenologically distinct processes are supported by distinct neural mechanisms. We hypothesised that inner speaking relies more strongly on an online motor-to-sensory simulation that constructs a multisensory experience, whereas inner hearing relies more strongly on a memory-retrieval process, where the multisensory experience is reconstructed from stored motor-to-sensory associations. Accordingly, we predicted that the speech motor system will be involved more strongly during inner speaking than inner hearing. This would be revealed by modulations of TMS evoked responses at muscle level following stimulation of the lip primary motor cortex. Overall, data collected from 31 participants corroborated this prediction, showing that inner speaking increases the excitability of the primary motor cortex more than inner hearing. Moreover, this effect was more pronounced during the inner production of a syllable that strongly recruits the lips (vs. a syllable that recruits the lips to a lesser extent). These results are compatible with models assuming that the primary motor cortex is involved during inner speech and contribute to clarify the neural implementation of the fundamental ability of silently speaking in one's mind.

© 2023 Published by Elsevier Ltd.

[☆] This is the preprint of a Stage-2 registered report. The Stage-1 version of this manuscript received in-principle acceptance at Cortex and is available at <https://osf.io/zcujn>.

* Corresponding author. Laboratoire de Psychologie Cognitive, AMU & CNRS (UMR 7290), France.

E-mail address: ladislav.nalborczyk@gmail.com (L. Nalborczyk).

<https://doi.org/10.1016/j.cortex.2023.09.007>

0010-9452/© 2023 Published by Elsevier Ltd.

1. Introduction

Rehearsing a conversation, preparing a public talk, or reading a novel are routine mental activities that are usually accompanied by an inner voice (Levine et al., 1982; Morin, 2012; Perrone-Bertolotti et al., 2012; Petkov & Belin, 2013; Sokolov, 1972). This inner voice feels like speech was produced or heard internally (e.g., Hurlburt & Heavey, 2015). In other words, it involves a conscious multisensory (e.g., auditory, kinaesthetic) experience (for reviews, see Alderson-Day & Fernyhough, 2015; Perrone-Bertolotti et al., 2014). Although commonly described as a unitary construct, inner speech may be better defined as a collection of distinct but related phenomenological experiences (e.g., Hurlburt, 2011; Hurlburt & Akhter, 2006). According to Hurlburt (2011), it is possible to distinguish the phenomenon of *inner speaking* from the phenomenon of *inner hearing*, whose feelings would be similar to talking in a tape recorder and hearing one's voice played back, respectively (Hurlburt et al., 2013).¹ The present research aims at better understanding the origins of the voice we experience during inner speaking and inner hearing, that is, the neurocognitive mechanisms through which it arises.

Stemming upon classical models of motor control, Lævenbruck et al. (2018) presented a predictive model of *wilful (expanded) inner speech* production (i.e., inner speaking). In this model, the auditory and kinaesthetic sensations perceived during inner speech are thought to be the predicted sensory consequences of (a copy of) inhibited speech motor acts. More precisely, these percepts are simulated by internal forward models that use the efference copies issued from an inverse model. According to this view, the primary motor cortex would be involved during inner speech production, but its output would be inhibited by prefrontal regions involved in response inhibition, such as the right inferior frontal cortex or the pre-supplementary motor area (Nalborczyk et al., 2022). The model put forward by Lævenbruck et al. (2018) predicts that the primary motor cortex will be involved during inner speaking, but it leaves unspecified the neural mechanisms supporting inner hearing and whether they differ from those supporting inner speaking.

According to the dual stream prediction model (Tian et al., 2016; Tian & Poeppel, 2012, 2013), the sensory content of inner speech is provided by two distinct processes. First, the sensory content of inner speech may result from a *motor simulation prediction stream*. In this view, inner speech would involve the same mechanisms as overt speech production except that the speech acts should be inhibited rather than executed (this proposal is similar to the model of Lævenbruck et al., 2018, although see below differences regarding the involvement of the primary motor cortex). Second, the sensory content of inner speech may be provided by an associative memory-based process called the *memory-retrieval prediction stream*

(Kosslyn et al., 1979; Moulton & Kosslyn, 2009; Tian & Poeppel, 2012). In this view, sensory percepts are motor-to-sensory associations established during past events and directly retrieved from long-term memory.²

The distinction between the motor simulation and memory retrieval streams has been linked to the distinction between inner speaking and inner hearing (e.g., Tian et al., 2018; Tian & Poeppel, 2012, 2013). Using fMRI, Tian et al. (2016) examined the neural correlates of articulation imagery (inner speaking) and hearing imagery (inner hearing) and observed that inner speaking more strongly recruits brain areas belonging to the motor-estimation stream whereas inner hearing more strongly recruits brain areas belonging to the memory-retrieval prediction stream. Moreover, these two imagery modes have been shown to have distinct MEG correlates and distinct modulatory effects on a subsequent /ba/-/da/ auditory categorisation task (Ma & Tian, 2019).

Critically, in the dual stream prediction model, the primary motor cortex is considered to be “bypassed” when producing inner speech (Tian et al., 2016; Tian & Poeppel, 2012, 2013). In comparison, the model introduced in Lævenbruck et al. (2018) and Grandchamp et al. (2019) predicts that the primary motor cortex is involved during inner speech production. In essence, at least some forms of inner speech may be accompanied by the emission of motor commands that are subsequently inhibited by cortical and subcortical mechanisms.

The hypothesis that inner speech involves motor inhibition is compatible with behavioural, lesional, and neuroimaging studies of motor imagery (for a review, see Guillot et al., 2012). The involvement of the primary motor cortex during inner speech and its partial inhibition by cortical and subcortical mechanisms may explain the residual peripheral muscular activity that is sometimes observed during inner speech production (Jeannerod, 2006; Lævenbruck et al., 2018; Nalborczyk, 2019). Another example is the observation that inner speech is accompanied by an increase in tongue motor excitability, compared to rest or to an auditory speech perception condition (Maegherman et al., 2020). If the dual stream prediction model is correct in that the primary motor cortex is “bypassed” during inner speech, neither inner speaking nor inner hearing should increase the involvement of the primary motor cortex. In contrast, if the motor control view is correct in that the primary motor cortex is involved (but actively inhibited by cortical and subcortical mechanisms), inner speaking should be accompanied by an increase in motor cortex excitability.

Transcranial magnetic stimulation (TMS) has been used to probe the involvement of the motor system in speech production and perception (for a review, see Devlin & Watkins, 2007). Single TMS pulses can be applied to the

¹ The distinction between inner speaking and inner hearing echoes previous distinctions such as the one between the *generative component* (i.e., the feeling of producing speech) and the *auditory component* (i.e., the feeling of hearing speech) of inner speech (e.g., MacKay, 1992) and the distinction between the *inner voice* and the *inner ear* in working memory (e.g., Baddeley et al., 1984; Buchsbaum & D'Esposito, 2019).

² The distinction between these two prediction streams is reminiscent of the distinction between the prediction-by-simulation and the prediction-by-association mechanisms in speech production and perception (Pickering & Garrod, 2013) and was previously discussed in more depth in Li et al. (2020), Ma and Tian (2019), Nalborczyk (2019), Nalborczyk et al. (2021), Tian and Poeppel (2012), and Tian and Poeppel (2013).

primary motor cortex to elicit motor evoked potentials (MEPs) in the orofacial muscles. MEPs can be recorded using surface electromyography (EMG). Their amplitude depends on the state of the motor system: it is greater during muscle contraction compared to rest. Therefore, MEP amplitude provides a direct measure of motor excitability that can be used to examine the involvement of the motor system during speech production and perception (Möttönen et al., 2014). For instance, seeing or hearing speech increases motor excitability in a content- and effector-specific manner, in the same way that observing hand movements does (e.g., Fadiga et al., 2002; Watkins et al., 2003). Sub-cortically, however, MEPs recorded from the orofacial muscles and those recorded from the hand muscles originate from different motor pathways. Whereas muscles from the face are innervated by the corticobulbar pathway, muscles from the hand are innervated by the corticospinal pathway, with direct consequences on the shape and latency of MEPs. Indeed, MEPs recorded over the orofacial muscles peak around 10–15 ms after the pulse, whereas MEPs recorded over the hand muscles peak around 20–25 ms after the pulse, because (amongst other things) of the different lengths of the corticobulbar vs. corticospinal tracts (Adank et al., 2018; Maegherman et al., 2020; Möttönen et al., 2014).

We hypothesise that during inner speaking, sensory (e.g., auditory) percepts are mostly provided by a motor-to-sensory simulation, whereas during inner hearing, sensory percepts are mostly reconstructed from stored perceptual information. Therefore, inner speaking should be accompanied by an increase in motor cortex excitability evidenced by larger MEPs recorded over orofacial muscles. Conversely, inner hearing should be accompanied by a much reduced or even absent increase in motor cortex excitability (**hypothesis #1**). Our secondary hypothesis is that this increase in motor cortex excitability is content-specific, as it is usually observed in TMS studies of speech perception (e.g., Fadiga et al., 2002; Watkins et al., 2003). We predict that lip motor cortex excitability increases should be more pronounced when covertly speaking a syllable that should strongly recruit the lips (i.e., /bu/, hereafter referred to as a “rounded” syllable) than when covertly speaking a syllable that should recruit the lips to a lesser extent (i.e., /gi/, hereafter referred to as a “spread” syllable) (**hypothesis #2**).

2. Methods

In the *Methods* and *Data analysis* sections, we report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study (Simmons et al., 2012). A pre-registered version of our protocol can be found on OSF: <https://osf.io/7kwv6/>.

2.1. Ethics information

The present research complies with the French national ethics regulation (agreement of CPP Sud Méditerranée I, ANSM national number 2017-A03614-49) and the Declaration of Helsinki. All participants provided informed consent and received a monetary compensation of 40€.

2.2. Design

The experimental design was fully within participants, with two crossed two-level factors: inner speech mode (inner speaking vs. inner hearing) and articulatory features (rounded vs. spread syllables). The experiment also included control trials during which we applied single-pulse TMS to assess motor cortex excitability while participants were performing an unrelated task. In these control trials, participants were asked to imagine tapping with their ipsilateral foot (i.e., to perform motor imagery of the foot; Fig. 1).

2.3. Procedure

2.3.1. Transcranial magnetic stimulation

Participants were familiarised with TMS before the training session. We asked the participants to fill out a safety screening form. Participants with contraindications for TMS (e.g., neurological disorders, medication, a family history of epilepsy) were not included in the experiment. Afterwards, we explained the details of the experiment to the participant and obtained their written informed consent.

To ensure similar levels of muscle activation throughout the experiment, we trained the participants to maintain a baseline activity between 20 and 30% of their maximum voluntary contraction. To find the location of the motor lip representation on the contralateral primary motor cortex, we followed the protocol described in Möttönen et al. (2014). We first localised the motor hand representation with the hotspot method, that is, by looking for the cortical site that elicits the maximal MEPs at a given intensity. Then, we localised the lip hotspot from the hand one, by keeping a minimal 5-sec break between TMS pulses.

The active motor threshold was defined as the minimal intensity necessary to elicit a lip MEP with a minimum peak-to-peak amplitude of approximately 50 μ V (as in Watkins et al., 2003; Watkins & Paus, 2004) in five stimulations out of ten (Rothwell et al., 1999). The intensity of the stimulator was then set to approximately 120% of the active motor threshold during the experimental session (this threshold was adjusted on a per-participant basis to maintain comfort throughout the experiment, cf. supplementary materials). As in Maegherman et al. (2020), we used a figure-of-eight coil placed around a 45° angle relative to the sagittal plane, inducing a posterior-to-anterior current flow, approximately perpendicular to the lateral fissure. The position of the TMS coil relative to cortex was continuously tracked and maintained throughout the experiment using a neuronavigation system (Navigation Brain System, Nexstim, Helsinki, Finland). A standard MRI image was used for the neuronavigation. All TMS pulses were monophasic and generated by a Magstim 200 device. The average active motor threshold was of 59.87% (ranging from 46% to 74%, SD = 7.83) (expressed as a percentage of maximum stimulator output).

2.3.2. Surface electromyography

EMG activity of the (right section of the) orbicularis oris (OO) muscle and the abductor pollicis brevis (APB) muscle of the right hand was recorded bipolarly with surface electrodes connected to a 16-channel amplifier (BrainAmp ExG with eight

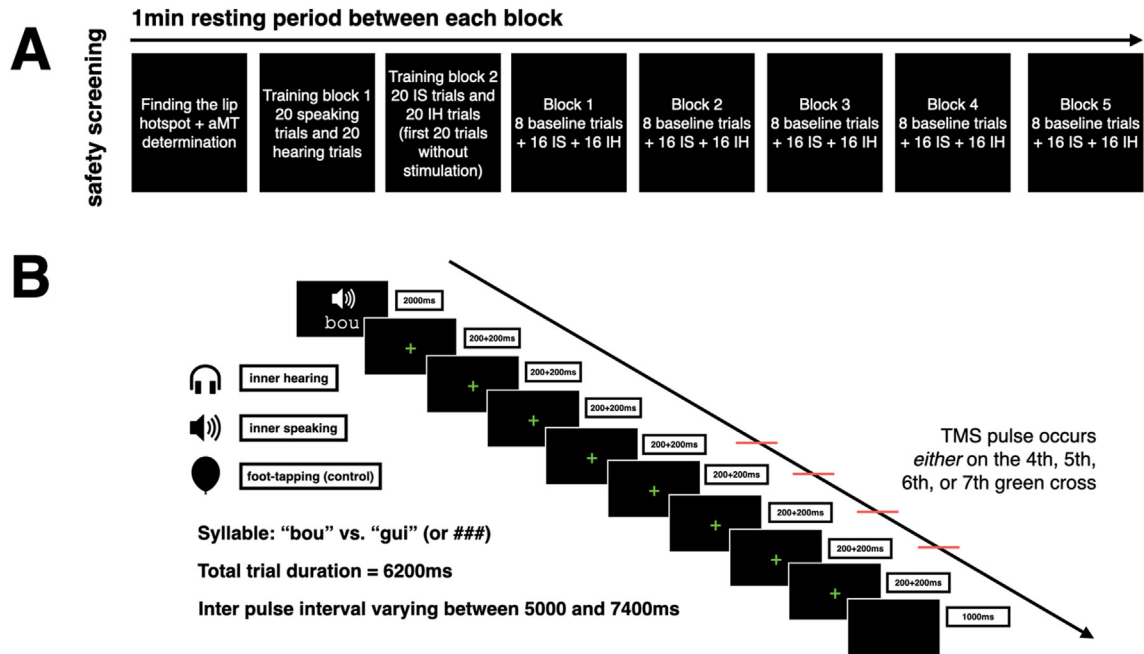


Fig. 1 – A. Experimental procedure. The main experimental part (post-training) involved 5 blocks of 40 trials each. aMT: active motor threshold, IS: inner speaking, IH: inner hearing. B. Illustrated timecourse of a single trial.

bipolar and eight auxiliary channels, Brain Products Company, Gilching, Germany) at a sampling rate of 5 kHz. Using alcohol, we cleaned the skin on the right section of the OO, above the belly of the APB muscle, as well as the skin around the site of the ground electrode placed on the right temple. We then attached electrodes on these sites. We visually checked the recorded EMG signals. If the signals appeared to be noisy, we re-cleaned the skin and re-attached the electrodes. After data collection, we computed the area-under-the-curve (AUC) of the rectified EMG signal. The covered area was taken as an index of the MEP size (Maegherman et al., 2020). In each trial, the MEP AUC was computed in a window spanning from 8 to 35 ms after the TMS pulse. As in Maegherman et al. (2020), a pre-pulse section of 27 ms (i.e., from 35 to 8 ms before the pulse) was extracted to allow a posteriori checks of equivalent baseline contraction across conditions.

2.3.3. Training blocks

Training was composed of two blocks. In the first block, participants were trained to either produce (overtly) or to listen to a syllable for a 2-sec period at a 2.5Hz pace (20 hearing and 20 speaking trials; Fig. 1). First, a syllable was presented on screen for 1 s. In speaking trials, participants had to repeatedly utter this syllable at 2.5Hz for 2 s. The 2.5Hz pace was cued by the display of a green cross on the screen. In hearing trials, participants were asked to listen to recordings of two native French speakers uttering the target syllables at a 2.5Hz pace. The syllables were chosen to induce a stronger involvement of the orbicularis oris muscle or of the zygomaticus muscle (i.e., /bu/vs./gi/, respectively). At the end of this first training block, participants were given earplugs to reduce the discomfort caused by the TMS click sound (Counter et al., 1991).

In the second training block, participants were familiarised with the phenomenological contrast between inner speaking and inner hearing conditions. In the inner speaking conditions, participants had to imagine speaking the syllables “in their mind” without moving the speech effector and without producing any sound (as in Tian et al., 2016). In the inner hearing conditions, participants had to recreate in their minds the voice from the hearing trials used in the first training block, while also minimising the feeling of movement in their speech effectors (as in Tian et al., 2016).

It should be noted that Tian and Poeppel (2012), Tian and Poeppel (2013), and Tian et al. (2016) use different definitions of inner speaking and inner hearing than Hurlburt et al. (2013). For the former, inner speaking refers to the act of silently (mentally) talking to oneself from the first perspective, with one's own perceived voice, whereas inner hearing refers to the act of imagining hearing speech, produced with the voice of someone else (sometimes designated as auditory verbal imagery). When defined in these terms, the distinction between inner speaking and inner hearing may be considered as the “speech analogue” of the distinction between first-person and third-person motor imagery in the motor imagery literature. As noted by Alderson-Day and Fernyhough (2015), however, this operationalisation of the phenomena of inner speaking and inner hearing makes it difficult to distinguish between the influence of the perspective and the distinction between inner speaking and inner hearing per se. While we acknowledge this limitation, we decided to use this operationalisation of inner speaking and inner hearing in order to compare the results of the present study to those of previous studies (e.g., Tian et al., 2016; Tian & Poeppel, 2012, 2013).

2.3.4. Experimental protocol

The timing of trials was identical across conditions (Fig. 1). In each trial, the target syllable (i.e., /bu/ vs. /gi/, written as “bou” vs. “gui”), together with a pictogram indicating whether the participant should generate inner speaking, inner hearing, or foot-tapping imagery, was displayed for 2000 ms. We used a pictogram that does not directly refer to the effector targeted by the TMS pulse to avoid automatic activation of the corresponding effector (see Fig. 1). Then, a green cross appeared at the center of the screen every 400 ms (i.e., at a 2.5 Hz pace) for 3200 ms. Participants were instructed to start producing the syllable at the appearance of the first green cross, and to keep repeating it in rhythm with the appearance of the green cross (i.e., every 400 ms). In all three task conditions, the TMS pulse exactly occurred exactly on the fourth, fifth, sixth, or seventh green cross. We decided to vary the occurrence of the TMS pulse to maintain the participant's attention on the task throughout the experiment and to reduce the predictability of the pulse's occurrence. At the end of each trial, the last (i.e., eighth) green cross was followed by a blank screen presented for 1000 ms. This procedure resulted in inter-pulse-intervals varying between 5000 ms and 7400 ms.

The main experimental part (post-training) consisted of five blocks of 40 trials each, yielding a total of 200 trials/MEPs per participant (40 control trials, 80 inner speaking trials: 40 /bu/ and 40 /gi/, and 80 inner hearing trials: 40 /bu/ and 40 /gi/). The order of trials within each experimental block was randomised across participants. Each block was followed by a 1 min resting period. Finally, participants had to fill out the Movement Imagery Questionnaire-3 Second French version (Robin et al., 2020). The experimental procedure was developed using the PsychoPy software (Peirce et al., 2019). Participants were then fully informed about the theoretical rationale for the study and compensated for their participation.

2.4. Sampling plan

To define the number of participants, we conducted a Bayesian a priori power analysis, where “statistical power” is to be understood in its general meaning, that is, the probability of achieving some statistical goal (Kruschke, 2015). We simulated data (see the supplementary materials for more details) by varying the number of participants (from 20 to 50) and the number of trials in each condition per participant (30 vs. 60). We were interested in the probability of detecting two effects: i) the main effect of the inner speech mode, that is, the difference between the average MEPs in the inner speaking vs. inner hearing conditions (60 trials in each condition), and ii) the effect of the type of syllable, that is, the difference between the average MEPs for /bu/ vs. /gi/ trials within the inner speaking condition (30 trials in each condition).

The null hypothesis (i.e., no difference between conditions) requires more observations to be corroborated than the alternative hypotheses of small, medium, or large effects (e.g., Schönbrodt et al., 2017; Schönbrodt & Wagenmakers, 2018). Therefore, we decided to plan for a null effect to obtain a conservative estimate of the number of observations and participants needed to detect all sorts of effects, from null to large. This analysis revealed that, with 30 or more observations per participant and per condition, we needed at least 30

participants to reach a probability equal or superior to 0.9 of obtaining a Bayes factor (BF) equal or superior to 10 in favour of the null hypothesis. The detailed resulting power curve and the reproducible code used to conduct this analysis are available in the online supplementary materials.

2.5. Participants

In accordance with our power analysis, we recruited 31 French-speaking undergraduate students in Psychology from Aix-Marseille University, ranging in age from 18 to 27 years ($M = 20.42$, $SD = 1.85$, 29F, 2M), with no reported history of psychiatric or neurological disorder, speech disorder, or hearing deficit.

2.6. Analysis plan

All analyses were conducted in R (R Core Team, 2021). We fitted several Bayesian multilevel models using the brms package (Bürkner, 2017) with two categorical predictors (recoded using sum contrasts as -0.5 vs. 0.5): the inner speech mode (i.e., inner speaking vs. inner hearing), the syllable (i.e., /bu/ vs. /gi/), and the AUC of the lip MEP as a continuous dependent variable (for an introduction to Bayesian multilevel modelling, see Nalborczyk et al., 2019). For each effect of interest, we report the mean of the posterior distribution along with its 95% credible interval, as well as the mean of the posterior distribution of the standardised mean difference (Cohen's d) with its 95% credible interval, when appropriate (i.e., when comparing two conditions). We also report the Bayes factor (BF), which quantifies the relative support (evidence) for either the null or the alternative hypothesis (e.g., Wagenmakers, 2007). Data collection and analysis were not performed blind to the conditions of the experiments.

We tested for differences in baseline background contraction across conditions per participant. If this test revealed a difference (i.e., a $BF_{10} > 10$), we applied the common range correction (as in Spieser et al., 2013; method adapted from Schieppati et al., 1996). This method consists in removing trials below the maximum value of minimum values across conditions and trials above the minimum value of maximum values across conditions, per participant. After removing these trials, we then tested again for a difference. If there still was a difference in baseline across conditions, we removed this participant from the subsequent analyses. If there was no difference across conditions, and if there were at least 20 trials per condition left for this participant, we kept it in the subsequent analyses. If the common range procedure removed more than 20 trials per condition and participant, we applied another procedure in which all trials are kept, consisting in including the background contraction level as a continuous predictor in the statistical models.

It should be noted that we finally opted for the latter approach, which gave similar results as the former one, with the advantage of not discarding the data crossing an arbitrary threshold, and therefore resulting in more precise estimates. More precisely, we added the (per-participant) level of background EMG activity (its standardised AUC) as both a fixed and a random effect in the model (cf. the model formula on page 5 of the supplementary materials, reproduced in a simplified

format below) (the detailed code is also available in the Rmarkdown file of the manuscript): `post ~ 1 + mode * syllable + pre + (1 + mode * syllable + pre | participant)`. Given this formulation, estimates of the model should be interpreted at the per-participant average level of background EMG contraction (i.e., the pre variable). In other words, estimates from this model can be interpreted in the following way: “Given (conditionally on) the level of background EMG contraction, what is the (additional/remaining) effect of mode, syllable, and their interaction mode: syllable?”. For completeness, analyses performed using the common range procedure are also reported in the online supplementary materials.

The first MEP in each block was removed because it is usually much larger than those that follow. As noted by Möttönen et al. (2014), it is not always possible to record robust lip MEP when the lips are “relaxed”. Therefore, we pre-registered that we would report the number of participants in which the experiment could not be carried out (i.e., participants for which the stimulation was uncomfortable and participants in which we could not elicit MEPs). Fortunately, this did not happen and no participant was discarded from the analyses for this reason.

3. Results

This section is divided into two parts. First, we present results from confirmatory (preregistered) analyses, aiming to test the difference in MEP amplitude between inner speaking and inner hearing (hypothesis #1) and the difference between the /bu/ and /gi/ syllables in the inner speaking conditions (hypothesis #2). Second, we present results from exploratory (non-preregistered) analyses, including an assessment of inter-individual differences in the two effects of interest, analyses of the relation between these effects and self-reported motor imagery skills, and analyses of the cortical silent period.

As predicted, inner speaking was associated with larger MEPs than inner hearing (hypothesis #1), and the mental production of the/bu/syllable was associated with larger MEPs than the mental production of the/gi/syllable in the inner speaking condition (hypothesis #2). However, this syllable effect was not specific to inner speaking (i.e., there was no evidence for an interaction effect between inner speech mode and syllable). Using a model comparison approach, we further demonstrated that these two effects were in the same direction in all participants.

3.1. Confirmatory (preregistered) analyses

Before moving to the statistical results, we represent the distribution of standardised MEP amplitudes across conditions in Fig. 2. This figure shows that the MEP amplitude recorded in the two inner hearing conditions was at similar levels as those recorded in the control condition. Conversely, MEPs were larger in the two inner speaking conditions. Moreover, MEPs were larger for the /bu/ syllable than for the /gi/ syllable.

To estimate these effects while accounting for the skewness of the collected data (for more details, see the online supplementary materials), we fitted a multilevel Skew-Normal model.

The Skew-Normal distribution is a generalisation of the Normal distribution with three parameters ξ (xi), ω (omega), and α (alpha) for location, scale, and skewness (shape), respectively. Estimates from this model regarding the location parameter are reported in Table 1.

First, notice that the effect of background (i.e., pre-pulse) EMG activity (i.e., the pre variable) was strongly positive ($\beta = 0.799$, 95% CrI [0.695, 0.907], $BF_{10} = 10 \times 10^{17}$), indicating that, on average, higher levels of background EMG activity were associated with larger MEPs. This phenomenon is well known and stresses again the importance of including the level of background EMG activity in the model when estimating the effect of the other variables of interest (here, the effect of inner speech mode, the effect of the syllable, and their interaction).

Regarding hypothesis #1 (i.e., the difference between inner speaking and inner hearing), this analysis revealed that MEPs were larger in the inner speaking than in the inner hearing conditions ($\beta = 0.103$, 95% CrI [0.054, 0.151], $BF_{+} = 59999$) and larger for the/bu/than for the/gi/syllable ($\beta = 0.078$, 95% CrI [0.032, 0.124], $BF_{+} = 1845.154$).³ There was weak evidence in favour of a null interaction effect ($\beta = 0.033$, 95% CrI [-0.07, 0.136], $BF_{10} = 0.064$).

Regarding hypothesis #2 (i.e., the difference between the/bu/and/gi/syllables in the inner speaking conditions), a contrast analysis revealed that there was strong evidence for larger MEPs during inner speaking of the /bu/ syllable than during inner speaking of the /gi/ syllable ($\beta = 0.094$, 95% CrI [0.035, 0.154], $BF_{+} = 213.286$).

3.2. Exploratory (non preregistered) analyses

In this section, we report the results of exploratory (i.e., non-preregistered) analyses. With these analyses, we aimed at assessing i) the variability of the observed effects across participants, ii) the impact of self-reported motor imagery abilities, and iii) potential differences between conditions in the duration of the cortical silent period (CSP), taken as an index of intracortical inhibition.

3.2.1. Inter-individual differences

Although group-level effects were small, they were remarkably stable across participants. We followed a model comparison approach that incorporates various constraints into Bayesian multilevel models (Haaf & Rouder, 2017; Rouder & Haaf, 2019) to test whether the estimated (true) effects were in the same direction for all participants (see also Van Geert et al., 2022, for a recent application). More precisely, we compared the evidence for a model that does not place any constraints on the participants' true effect (hereafter the “unconstrained” model) with the evidence for a model that constrains true participants' effect to have a particular sign (hereafter the “positive effects” model). The Bayes factor comparing the likelihood of the observed data under these two models was 127.25 (inverse $BF = 0.008$) for the effect of inner speech mode, indicating that the observed data were

³ For one-sided hypotheses, BF_{+} represents the ratio of the posterior probability of the effect being positive and the posterior probability of the effect being negative.

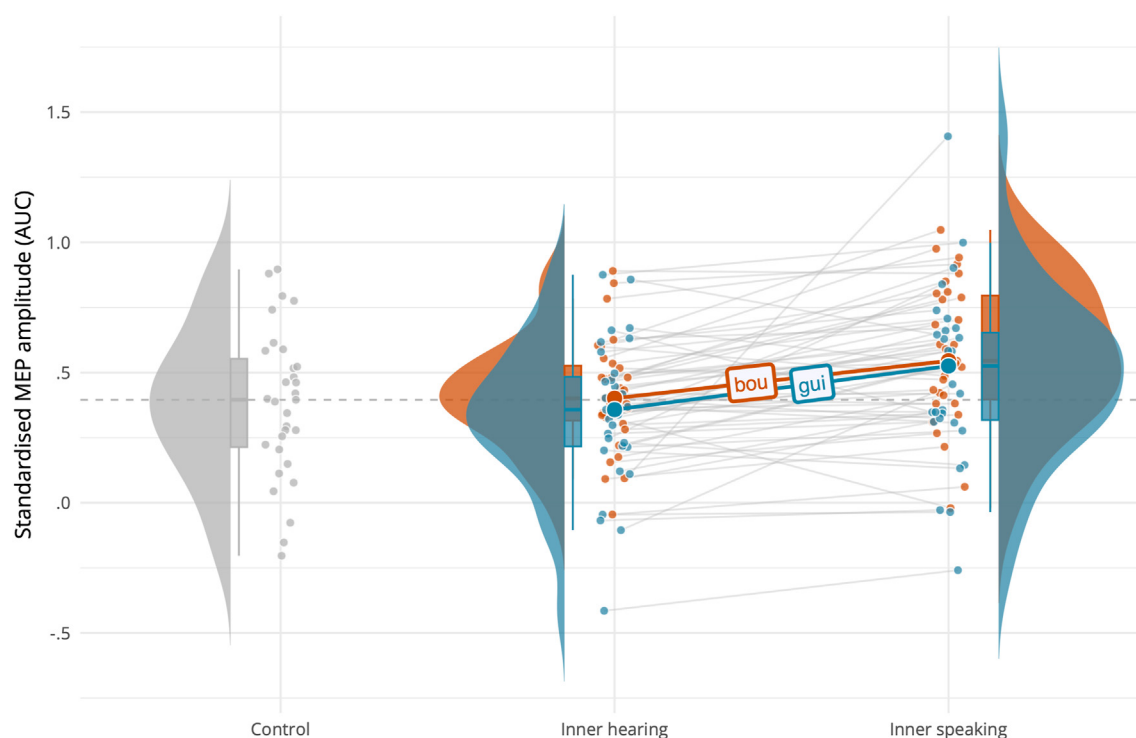


Fig. 2 – Average standardised MEP amplitude across conditions. Each dot represents a participant's median AUC computed across 40 trials per condition. The dashed grey horizontal line represents the group's median AUC in the control (imagined foot-tapping) condition.

127.25 more likely under the positive effects model than under the unconstrained model.

The Bayes factor comparing the likelihood of the observed data under the positive effects model and under the unconstrained model was 132.47 (inverse BF = 0.008) for the effect of the syllable, indicating that the observed data were 132.47 more likely under the positive effects model than under the unconstrained model. The Bayes factor comparing the likelihood of the observed data under the positive effects model and under the unconstrained model was 9.426 (inverse BF = 0.106) for the interaction effect, indicating that the observed data were 9.426 more likely under the positive effects model than under the unconstrained model.

Fig. 3 illustrates individual-level estimates of each effect based on the previously described Skew-Normal model.

Overall, these analyses suggest that *all* individuals show the same effects of inner speech mode and syllable: inner speaking led to larger MEPs than inner hearing, and mentally producing the /bu/ syllable led to larger MEPs than mentally producing the /gi/ syllable.

3.2.2. Impact of self-reported motor imagery abilities

To assess the impact of self-reported motor imagery abilities, we created a set of additional regression models containing either the score on each of the MIQ subscale (i.e., the internal perspective score, the external perspective score, or the kinesthetic score) or the total score. We then compared these models using the Widely Applicable Information Criterion (WAIC, Watanabe, 2010), a generalisation of the Akaike information criterion (Akaike, 1974). The WAIC provides a

Table 1 – Estimates from the multilevel Skew-Normal model regarding the location parameter.

Predictor	Estimate	SE	Lower	Upper	Rhat	BF ₁₀	BF ₊
mode	0.103	0.025	0.054	0.151	1.000	53.476	59999.000
syllable	0.078	0.024	0.032	0.124	1.000	5.645	1845.154
pre	0.799	0.054	0.695	0.907	1.000	10×10^{17}	∞
mode:syllable	0.033	0.052	-0.070	0.136	1.000	0.064	2.861

Note. The 'Estimate' column represents the estimated group-level effect (slope) of each predictor included in the model (in terms of standardised AUCs). The 'Lower' and 'Upper' columns contain the lower and upper bounds of the 95% CrI, whereas the 'Rhat' column reports the Gelman-Rubin statistic. The last two columns report the BF in favour of the alternative hypothesis (relative to the null) and the directional (i.e., one-sided) BF, respectively.

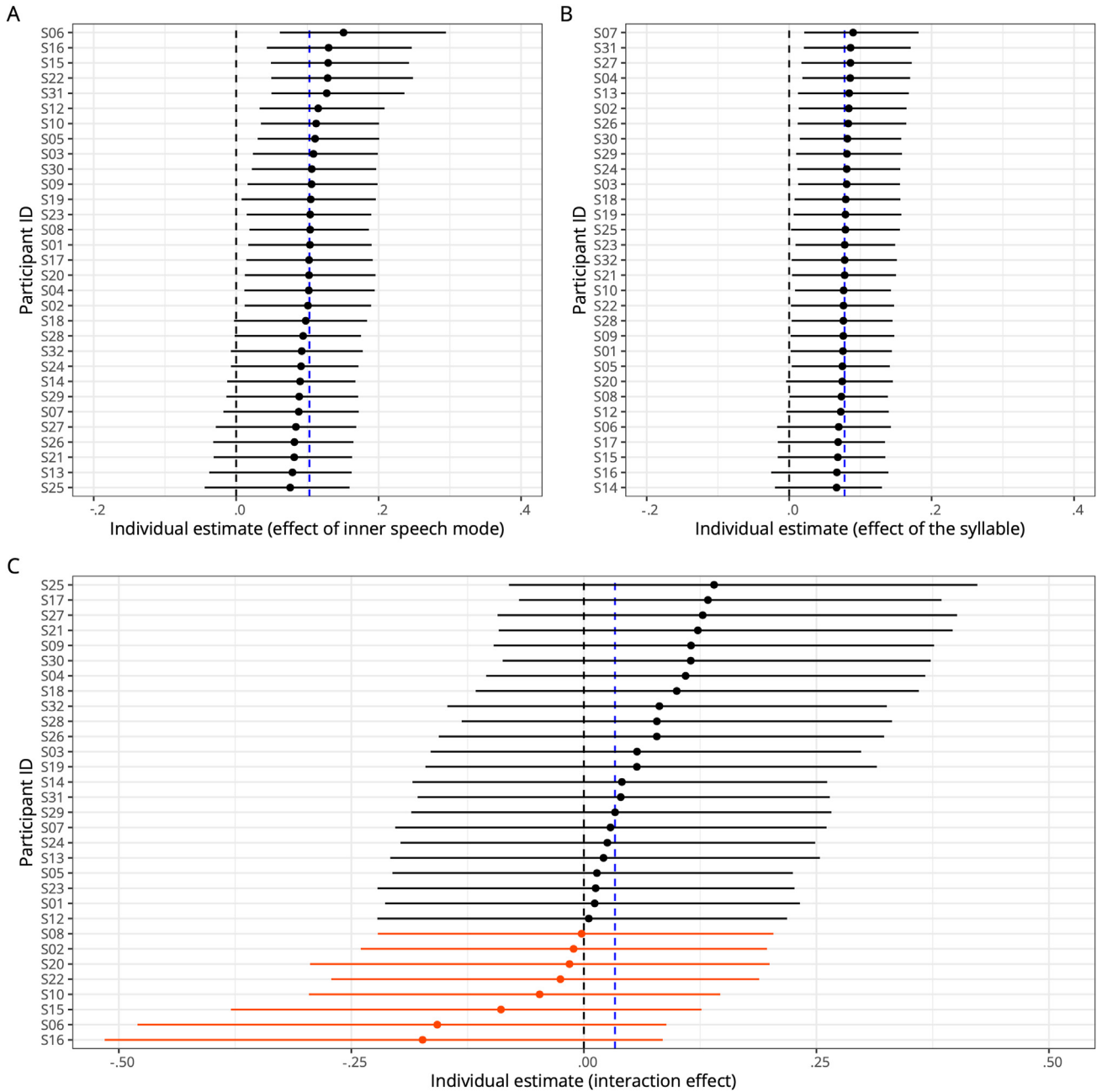


Fig. 3 – Model-based estimates of each individual-level (i.e., 'random' or 'varying') effects, in descending order. Negative estimates are highlighted in orange. The vertical blue dashed line represents the average effect. Panel A: Individual-level effects of the inner speech mode (positive values being associated with larger MEPs during inner speaking). Panel B: Individual-level effects of the syllable (positive values being associated with larger MEPs when producing the/bu/syllable). Panel C: Individual-level interaction effects between inner speech mode and syllable. Positive values are associated with a stronger syllable effect in the inner speaking condition (or symmetrically, a stronger effect of inner speech mode for the/bu/syllable).

relative measure of predictive accuracy of the models (it is an approximation of the out-of-sample deviance) and balances underfitting and overfitting by sanctioning models for their complexity (Burnham & Anderson, 2002; Burnham et al., 2011; Hegyi & Garamszegi, 2011). These analyses revealed that the

model with the lowest WAIC (i.e., the most parsimonious model) was the model without any MIQ score, suggesting that self-reported motor imagery abilities did not affect the previously described effects (see the online supplementary materials for code details).

Table 2 – Estimates from the multilevel Log-Normal model.

Predictor	Estimate	SE	Lower	Upper	Rhat	BF ₁₀
mode	0.203	0.195	−0.171	0.590	1.003	0.142
syllable	0.132	0.189	−0.232	0.515	1.003	0.099
mode:syllable	0.025	0.377	−0.714	0.767	1.003	0.172

Note. The 'Estimate' column represents the estimated effect (slope) of each predictor included in the model. The 'Lower' and 'Upper' columns contain the lower and upper bounds of the 95% CrI, whereas the 'Rhat' column reports the Gelman-Rubin statistic. The last column reports the Savage-Dickey BF in favour of the alternative hypothesis (relative to the null).

3.2.3. Analyses of the cortical silent period

The cortical silent period (CSP) refers to a period of TMS-induced reduction in the EMG activity of a voluntarily contracting muscle (for review, see [Hupfeld et al., 2020](#)). The duration of the CSP is obtained by measuring the time interval between the offset of the MEP and the restoration of EMG activity. Overall, the duration of the CSP is considered to reflect the levels of slow metabotropic postsynaptic GABA_A-mediated inhibition, occurring within the primary motor cortex ([Cardellicchio et al., 2020](#); [Hallett, 2007](#); [Moezzi et al., 2018](#); [Werhahn et al., 1999](#)). Crucially, intracortical inhibition has been suggested as one of the mechanisms preventing motor execution during motor imagery (for review, see [Guillot et al., 2012](#)).

To examine whether our different manipulations induced different levels of intracortical inhibition, we analysed the effect of inner speech mode (inner speaking vs. inner hearing) and the effect of the syllable to be produced mentally (/bu/ vs. /gi/) on the duration of the CSP (for more details on the determination of the CSPs' duration, see the online supplementary materials). To estimate these effects, we fitted a multilevel Log-Normal regression model to the CSPs' durations. Estimates from this model are reported in [Table 2](#). Overall, all effects were small and more likely to appear under the null hypothesis than under the alternative hypothesis, suggesting that levels of intracortical inhibition did not differ across conditions.

4. Discussion

We investigated the neural processes involved in two forms of inner speech: inner speaking and inner hearing. Based on previous observations and on predictions of the dual stream prediction model ([Tian et al., 2016](#); [Tian & Poeppel, 2012, 2013](#)), we assumed that inner speaking relies more strongly on a motor simulation mechanism than inner hearing. Based on the neurocognitive model of inner speech production developed in [Løvenbruck et al. \(2018\)](#) and [Grandchamp et al. \(2019\)](#), we hypothesised that the motor simulation mechanism underlying inner speaking would be indexed by increased levels of cortical excitability during inner speaking as compared to inner hearing (hypothesis #1). Given the involvement of motor simulation during inner speaking, we further hypothesised that the increase in cortical excitability during inner speaking should reflect the phonetic features of what is said (simulated) mentally (hypothesis #2).

Overall, these predictions were corroborated by our data. Even when controlling for pre-pulse EMG activity, we

observed that inner speaking was associated with larger MEPs than inner hearing (hypothesis #1), and that the mental production of the /bu/ syllable was associated with larger lip MEPs than the mental production of the /gi/ syllable in the inner speaking condition (hypothesis #2). However, this syllable effect was also present during inner hearing (it was only slightly stronger in the inner speaking condition, cf. [Table 1](#)), although the average MEP amplitude during inner hearing was not different from the average MEP amplitude in the control condition. We further observed that the two effects of interest pointed in the same direction for all participants ([Fig. 3](#)). These results provide explicit constraints for current models of inner speech production.

Our results are compatible with the distinction between inner speaking and inner hearing, as postulated by the dual stream prediction model. The stronger increase in cortical excitability during inner speaking than inner hearing supports the involvement of a motor simulation mechanism during inner speaking, whose role would be to provide the sensory content of inner speech (e.g., the inner voice). The stronger involvement of the primary motor cortex during inner speaking is consistent with the observation that inner speaking is associated with a stronger perceptual reactivation in auditory cortices ([Tian et al., 2016](#)). Whereas the syllable effect was stronger during inner speaking, it was nonetheless present (non-null) during inner hearing ($\beta = 0.061$, 95% CrI [0.005, 0.117], $BF_{+} = 25.625$). This result suggests that the motor simulation stream may also be solicited during inner hearing, but to a lesser extent than during inner speaking (consistent with the results obtained by [Tian et al., 2016](#)).⁴ By contrasting inner speaking and inner hearing of non-speech sounds, [Chu et al. \(2023\)](#) recently showed a clear dissociation in the involvement of the motor-based and memory-based networks, further supporting the distinction between these two processes. Overall, our results are also compatible with the framework recently provided by [Pratts et al. \(2023\)](#), in which inner speech can be generated by two separate mechanisms similar to those postulated by the dual stream prediction model, according to the intentionality and egocentricity constraints of the task.

The dual stream prediction model grants a secondary role to the primary motor cortex for inner speech, as it is considered to be “bypassed” during inner speech ([Tian et al., 2016](#); [Tian & Poeppel, 2012, 2013](#)). Our results contradict this view by

⁴ A more trivial but unverifiable interpretation of this result is that some trials contained inner speech performed in the incorrect mode. This would, as per our hypotheses, increase the average AUC in inner hearing trials, or decrease it in inner speaking trials, or both.

showing that the primary motor cortex is involved during inner speaking and that this involvement is modulated by phonetic features. In contrast, our results are compatible with the models developed in Lævenbruck et al. (2018) and Grandchamp et al. (2019), as well as other models of motor imagery, such as Grush (2004), in which the role of the primary motor cortex during inner speech is to issue motor commands from which predictions of sensory consequences can be subsequently computed (leading to the rich multisensory content of inner speech). However, the observed increase in M1 excitability could be due to various reasons, for example strategically performing the imagery tasks by subliminally executing the action. Moreover, our observations cannot exclude the possibilities of upper motor pathways mediating inner speaking. These questions can not be answered from the present data and could be targeted in subsequent studies. Nevertheless, the observed increase in M1 excitability revives a classic crucial issue referred to as “the problem of inhibition of execution” by Jeannerod (2001): Given the involvement of the motor system in providing the multisensory content of inner speech, how is it possible for inner speech not to lead to motor execution?

It has been suggested that the subthreshold involvement of the primary motor cortex may result from either a subliminal activation or from active inhibitory mechanisms counteracting a supraliminal activation. Crucially, both options require an explanation of how activity within the primary motor cortex is maintained under the execution threshold. Regarding the first mechanism, Bach et al. (2021) suggested that the motor (execution) threshold may be “upregulated” during motor imagery to prevent execution. How this would be achieved or implemented however, is not specified (Nalborczyk et al., 2023). Regarding the second mechanism, supraliminal activation of the motor system could be counterbalanced by parallel inhibitory signals (e.g., Berthoz, 1996; Bonnet et al., 1997; Jeannerod, 1994, 2001). Recent behavioural results obtained using an action-mode switching paradigm support the hypothesis of parallel inhibitory mechanisms operating during motor imagery (Rieger et al., 2017). By asking participants to rapidly alternate between imagined and executed movements, it is possible to measure switching costs or benefits when switching from imagery to execution or from execution to imagery. Overall, results from such studies show that motor imagery of hand movements slows down performance in the subsequent trial (Bart et al., 2021a; 2021b, 2021c; Rieger et al., 2017; Scheil & Liefoghe, 2018). Accordingly, we previously proposed that similar inhibitory mechanisms may also be at play during inner speech production to prevent the execution of speech acts (Grandchamp et al., 2019; Lævenbruck et al., 2018; Nalborczyk et al., 2022).

Speech production differs considerably from the simple hand movements which are often assessed in motor imagery studies. In particular, it requires the coordination and sequencing of many articulators in short timescales. Zhao et al. (2023) suggested that these peculiarities may require an additional (or alternative) cerebral network for inhibiting speech. Using high-density ECoG, they observed activity in the premotor cortex associated with speech stopping. Moreover, electrocortical stimulation over this area caused involuntary speech arrests, interpreted as an engagement of the inhibitory

mechanisms implemented within this area (see also Silva et al., 2022). This would be consistent with recent models of inner speech production, in which parallel inhibitory mechanisms are assumed to be issued by the rostral part of the precentral gyrus (Lævenbruck et al., 2018) or the orbitofrontal cortex (Grandchamp et al., 2019). Although our data cannot decide between these possibilities, they provide preliminary evidence regarding the role of intracortical inhibition during inner speech. Our analyses of the CSP durations revealed that levels of intracortical inhibition did not differ across inner speech modes (i.e., inner speaking vs. inner hearing) nor across syllables (i.e., bou vs. gui). This result goes against the hypothesis of an increased intracortical (GABA_b-mediated) inhibition during inner speaking. However, it should be stressed that it does not allow ruling out the involvement of other forms of inhibition, such as those involving GABA_a-mediated intracortical inhibition or cortico-subcortico-cortical circuits. Further research should aim at clarifying how these multiple processes interact together to maintain the activity of the primary motor cortex below the execution threshold during inner speech. Examining how the interplay between excitatory and inhibitory inputs to the primary motor cortex is modulated in different forms of inner speech (for instance in dysfunctional inner speech such as rumination or auditory verbal hallucinations) and their precise neural implementation are important future directions.

In summary, the results we describe establish the differential involvement of the primary motor cortex in two different phenomenological experiences of inner speech, suggesting that distinct neural processes can support the mental production of speech. Various forms of inner speech, such as inner speaking or inner hearing, selectively engage these processes, and their involvement can be probed using transcranial magnetic stimulation. These results stress the importance of examining different forms of inner speech to account for its variety.

Data availability

Data are available via the Open Science Framework: <https://osf.io/7kwv6/>.

Code availability

Supplementary materials are available via the Open Science Framework: <https://osf.io/7kwv6/>.

Author contributions

Conceptualisation: LN, ML, MB, LS, FXA; Data curation: LN, MB, VS, LS; Formal analysis: LN, LS; Funding acquisition: LN, ML, LS, FXA; Investigation: LN, LS; Methodology: LN, ML, MB, VS, LS, FXA; Project administration: LN, ML, MB, LS, FXA; Resources: LN, ML, MB, LS, FXA; Software: LN, LS, VS; Supervision: ML, MB, LS, FXA; Validation: ML, FXA; Visualisation: LN; Writing - original draft: LN; Writing - review and editing: LN, ML, MB, LS, FXA.

Open practices section

The study in this article earned Open Data, Open Material and Preregistered badges for transparent practices. The data, materials and preregistered studies are available at: <https://osf.io/7kwv6/>

Declaration of competing interest

The authors declare no competing interests.

Acknowledgements

We want to thank Camille Grasso and Kate Watkins for insightful comments at various stages of the present research, and Suong Nguyen for her help with setting up the stimulation protocol and the analysis scripts. This work, carried out within the Institute of Convergence ILCB (ANR-16-CONV-0002), has benefited from support from the French government (France 2030), managed by the French National Agency for Research (ANR) and the Excellence Initiative of Aix-Marseille University (A*MIDEX). The funders have/had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cortex.2023.09.007>.

REFERENCES

- Adank, P., Kennedy-Higgins, D., Maegherman, G., Hannah, R., & Nuttall, H. E. (2018). Effects of coil orientation on motor evoked potentials from orbicularis oris. *The Florida Nurse*, 12. <https://doi.org/10.3389/fmins.2018.00683>
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), Article 716723. <https://doi.org/10.1109/tac.1974.1100705>
- Alderson-Day, B., & Fernyhough, C. (2015). Inner speech: Development, cognitive functions, phenomenology, and neurobiology. *Psychological Bulletin*, 141(5), 931–965. <https://doi.org/10.1037/bul0000021>
- Bach, P., Frank, C., & Kunde, W. (2021). Why motor imagery isn't really motoric: Towards a reconceptualization in terms of effect-based action control. [Preprint]. *PsyArXiv*. <https://doi.org/10.31234/osf.io/bkf76>
- Baddeley, A. D., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *The Quarterly Journal of Experimental Psychology Section A*, 36(2), 233–252. <https://doi.org/10.1080/14640748408402157>
- Bart, V. K. E., Koch, I., & Rieger, M. (2021a). Decay of inhibition in motor imagery. *The Quarterly Journal of Experimental Psychology: QJEP*, 74(1), 77–94. <https://doi.org/10.1177/1747021820949388>
- Bart, V. K. E., Koch, I., & Rieger, M. (2021b). Inhibitory mechanisms in motor imagery: Disentangling different forms of inhibition using action mode switching. *Psychological Research*, 85(4), 1418–1438. <https://doi.org/10.1007/s00426-020-01327-y>
- Bart, V. K. E., Koch, I., & Rieger, M. (2021c). Expectations affect the contribution of tonic global inhibition, but not of phasic global inhibition to motor imagery. *Journal of Experimental Psychology: Human Perception and Performance*. <https://doi.org/10.1037/xhp0000961>
- Berthoz, A. (1996). The role of inhibition in the hierarchical gating of executed and imagined movements. *Cancer Biotherapy & Radiopharmaceuticals*, 3(2), 101–113. [https://doi.org/10.1016/0926-6410\(95\)00035-6](https://doi.org/10.1016/0926-6410(95)00035-6)
- Bonnet, M., Decety, J., Jeannerod, M., & Requin, J. (1997). Mental simulation of an action modulates the excitability of spinal reflex pathways in man. *Cancer Biotherapy & Radiopharmaceuticals*, 5(3), 221–228. [https://doi.org/10.1016/s0926-6410\(96\)00072-9](https://doi.org/10.1016/s0926-6410(96)00072-9)
- Buchsbaum, B. R., & D'Esposito, M. (2019). A sensorimotor view of verbal working memory. *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior*, 112, 134–148. <https://doi.org/10.1016/j.cortex.2018.11.010>
- Bürkner, P.-C. (2017). Brms : An r package for bayesian multilevel models using stan. *Journal of Statistical Software*, 80(1). <https://doi.org/10.18637/jss.v080.i01>
- Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd ed.). Springer.
- Burnham, K. P., Anderson, D. R., & Huyvaert, K. P. (2011). AIC model selection and multimodel inference in behavioral ecology: Some background, observations, and comparisons. *Biomedical and Environmental Sciences: BES*, 65(1), 23–35. <https://doi.org/10.1007/s00265-010-1029-6>
- Cardellicchio, P., Dolfini, E., Fadiga, L., & D'Ausilio, A. (2020). Parallel fast and slow motor inhibition processes in Joint Action coordination. *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior*, 133, 346–357. <https://doi.org/10.1016/j.cortex.2020.09.029>
- Chu, Q., Ma, O., Hang, Y., & Tian, X. (2023). Dual-stream cortical pathways mediate sensory prediction. *Cerebral Cortex*, 33(14), 8890–8903. <https://doi.org/10.1093/cercor/bhad168>
- Counter, S. A., Borg, E., & Lofqvist, L. (1991). Acoustic trauma in extracranial magnetic brain stimulation. *Electroencephalography and Clinical Neurophysiology*, 78(3), 173–184. [https://doi.org/10.1016/0013-4694\(91\)90031-X](https://doi.org/10.1016/0013-4694(91)90031-X)
- Devlin, J. T., & Watkins, K. E. (2007). Stimulating language: Insights from TMS. *Brain: a Journal of Neurology*, 130(3), 610–622. <https://doi.org/10.1093/brain/awl331>
- Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience*, 15(2), 399–402. <https://doi.org/10.1046/j.0953-816x.2001.01874.x>
- Grandchamp, R., Rapin, L., Perrone-Bertolotti, M., Pichat, C., Haldin, C., Cousin, E., Lachaux, J.-P., Dohen, M., Perrier, P., Garnier, M., Baciu, M., & Levenbruck, H. (2019). The ConDialInt model: Condensation, dialogality, and intentionality dimensions of inner speech within a hierarchical predictive control framework. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.02019>
- 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>
- Guillot, A., Di Rienzo, F., MacIntyre, T., Moran, A., & Collet, C. (2012). Imagining is not doing but involves specific motor commands: A review of experimental data related to motor inhibition. *Frontiers in Human Neuroscience*, 6. <https://doi.org/10.3389/fnhum.2012.00247>

- Haaf, J. M., & Rouder, J. N. (2017). Developing constraint in Bayesian mixed models. *Psychological Methods*, 22(4), 779–798. <https://doi.org/10.1037/met0000156>
- Hallett, M. (2007). Transcranial magnetic stimulation: A primer. *Neuron*, 55(2), 187–199. <https://doi.org/10.1016/j.neuron.2007.06.026>
- Hegyí, G., & Garamszegi, L. Z. (2011). Using information theory as a substitute for stepwise regression in ecology and behavior. *Biomedical and Environmental Sciences: BES*, 65(1), 69–76. <https://doi.org/10.1007/s00265-010-1036-7>
- Hupfeld, K. E., Swanson, C. W., Fling, B. W., & Seidler, R. D. (2020). TMS-induced silent periods: A review of methods and call for consistency. *Journal of Neuroscience Methods*, 346, Article 108950. <https://doi.org/10.1016/j.jneumeth.2020.108950>
- Hurlburt, R. T. (2011). *Investigating pristine inner experience: Moments of truth*. C. U. Press.
- Hurlburt, R. T., & Akhter, S. A. (2006). The descriptive experience sampling method. *Phenomenology and the Cognitive Sciences*, 5(3–4), 271–301. <https://doi.org/10.1007/s11097-006-9024-0>
- Hurlburt, R. T., & Heavey, C. L. (2015). Investigating pristine inner experience: Implications for experience sampling and questionnaires. *Consciousness and Cognition*, 31, 148–159. <https://doi.org/10.1016/j.concog.2014.11.002>
- Hurlburt, R. T., Heavey, C. L., & Kelsey, J. M. (2013). Toward a phenomenology of inner speaking. *Consciousness and Cognition*, 22(4), 1477–1494. <https://doi.org/10.1016/j.concog.2013.10.003>
- Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, 17(2), 187. <https://doi.org/10.1017/S0140525X00034026>
- Jeannerod, M. (2001). Neural simulation of action: A unifying mechanism for motor cognition. *Neuroimage*, 14(1), S103–S109. <https://doi.org/10.1006/nimg.2001.0832>
- Jeannerod, M. (2006). *Motor cognition: What actions tell the self*. Oxford University Press.
- Kosslyn, S. M., Pinker, S., Smith, G. E., & Schwartz, S. P. (1979). On the demystification of mental imagery. *Behavioral and Brain Sciences*, 2(4), 535–548. <https://doi.org/10.1017/S0140525X00064268>
- Kruschke, J. K. (2015). *Doing bayesian data analysis: A tutorial with R, JAGS, and stan* (2nd ed.). Academic Press.
- Lævenbruck, H., Grandchamp, R., Rapin, L., Nalborczyk, L., Dohen, M., Perrier, P., Baciú, M., & Perrone-Bertolotti, M. (2018). A cognitive neuroscience view of inner language: To predict and to hear, see, feel. In P. Langland-Hassan, & A. Vicente (Eds.), *Inner speech: New voices* (p. 37). Oxford University Press.
- Levine, D. N., Calvanio, R., & Popovics, A. (1982). Language in the absence of inner speech. *Neuropsychologia*, 20(4), 391–409. [https://doi.org/10.1016/0028-3932\(82\)90039-2](https://doi.org/10.1016/0028-3932(82)90039-2)
- Li, S., Zhu, H., & Tian, X. (2020). Corollary discharge versus efference copy: Distinct neural signals in speech preparation differentially modulate auditory responses. *Cerebral Cortex*, 30(11), 5806–5820. <https://doi.org/10.1093/cercor/bhaa154>
- MacKay, D. G. (1992). Constraints on theories of inner speech. In D. Reisberg (Ed.), *Auditory imagery* (pp. 121–149). Hillsdale, N. J: Erlbaum.
- Maegherman, G., Nuttall, H., Devlin, J., & Adank, P. (2020). The involvement of articulatory motor cortex in motor imagery of speech. <https://doi.org/10.31219/osf.io/5k9wa>
- Ma, O., & Tian, X. (2019). Distinct mechanisms of imagery differentially influence speech perception. *eNeuro*. <https://doi.org/10.1523/ENEURO.0261-19.2019>
- Moezzi, B., Schaworonkow, N., Plogmacher, L., Goldsworthy, M. R., Hordacre, B., McDonnell, M. D., Iannella, N., Ridding, M. C., & Triesch, J. (2018). Simulation of electromyographic recordings following transcranial magnetic stimulation. *Journal of Neurophysiology*, 120(5), 2532–2541. <https://doi.org/10.1152/jn.00626.2017>
- Morin, A. (2012). Inner speech. In V. S. Ramachandran (Ed.), *Encyclopedia of human behavior* (Second Edition, pp. 436–443). Academic Press. <https://doi.org/10.1016/B978-0-12-375000-6.00206-8>
- Möttönen, R., Rogers, J. C., & Watkins, K. E. (2014). Stimulating the lip motor cortex with transcranial magnetic stimulation. *Journal of Visualized Experiments : JoVE*, 88, 1–7. <https://doi.org/10.3791/51665>
- Moulton, S. T., & Kosslyn, S. M. (2009). Imagining predictions: Mental imagery as mental emulation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1521), 1273–1280. <https://doi.org/10.1098/rstb.2008.0314>
- Nalborczyk, L. (2019). Understanding rumination as a form of inner speech: Probing the role of motor processes [PhD] [thesis]. Univ. Grenoble Alpes & Ghent University. <https://thesiscommons.org/p6dct/>
- Nalborczyk, L., Banjac, S., Baeyens, C., Grandchamp, R., Koster, E. H. W., Perrone-Bertolotti, M., & Lævenbruck, H. (2021). Dissociating facial electromyographic correlates of visual and verbal induced rumination. *International Journal of Psychophysiology*, 159, 23–36. <https://doi.org/10.1016/j.ijpsycho.2020.10.009>
- Nalborczyk, L., Batailler, C., Lævenbruck, H., Vilain, A., & Bürkner, P.-C. (2019). An introduction to bayesian multilevel models using brms: A case study of gender effects on vowel variability in standard Indonesian. *Journal of Speech, Language, and Hearing Research: JSLHR*, 62(5), 1225–1242. https://doi.org/10.1044/2018_JSLHR-S-18-0006
- Nalborczyk, L., Debarnot, U., Longcamp, M., Guillot, A., & Alario, F.-X. (2022). The role of motor inhibition during covert speech production. *Frontiers in Human Neuroscience*, 16, Article 804832. <https://doi.org/10.3389/fnhum.2022.804832>
- Nalborczyk, L., Longcamp, M., Gajdos, T., Servant, M., & Alario, F.-X. (2023). Towards formal models of inhibitory mechanisms involved in motor imagery: A commentary on bach, 2022. frank, & kunde. <https://doi.org/10.31234/osf.io/tz6x2>
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1), 195–203. <https://doi.org/10.3758/s13428-018-01193-y>
- Perrone-Bertolotti, M., Kujala, J., Vidal, J. R., Hamame, C. M., Ossandon, T., Bertrand, O., Minotti, L., Kahane, P., Jerbi, K., & Lachaux, J.-P. (2012). How silent is silent reading? Intracerebral evidence for top-down activation of temporal voice areas during reading. *Journal of Neuroscience*, 32(49), 17554–17562. <https://doi.org/10.1523/JNEUROSCI.2982-12.2012>
- Perrone-Bertolotti, M., Rapin, L., Lachaux, J. P., Baciú, M., & Lævenbruck, 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>
- Petkov, C. I., & Belin, P. (2013). Silent reading: Does the brain “hear” both speech and voices? *Current Biology*, 23(4), R155–R156. <https://doi.org/10.1016/j.cub.2013.01.002>
- Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36(4), 329–347. <https://doi.org/10.1017/S0140525X120001495>
- Pratts, J., Pobric, G., & Yao, B. (2023). Bridging phenomenology and neural mechanisms of inner speech: ALE meta-analysis on egocentricity and intentionality in a dual-mechanistic framework. <https://doi.org/10.31219/osf.io/gaum4>
- R Core Team. (2021). R: A language and environment for statistical computing [manual]. <https://www.R-project.org/>
- Rieger, M., Dahm, S. F., & Koch, I. (2017). Inhibition in motor imagery: A novel action mode switching paradigm.

- Psychonomic Bulletin & Review, 24(2), 459–466. <https://doi.org/10.3758/s13423-016-1095-5>
- Robin, N., Coudevylle, G. R., Guillot, A., & Toussaint, L. (2020). French translation and validation of the movement imagery questionnaire-third version (MIQ-3f). *Movement & Sport Sciences - Science & Motricité*. <https://doi.org/10.1051/sm/2019035>
- Rothwell, J. C., Hallett, M., Berardelli, A., Eisen, A., Rossini, P., & Paulus, W. (1999). Magnetic stimulation: Motor evoked potentials. *The international federation of clinical neurophysiology. Electroencephalography and Clinical Neurophysiology*, 52(Supplement), 97–103.
- Rouder, J. N., & Haaf, J. M. (2019). A psychometrics of individual differences in experimental tasks. *Psychonomic Bulletin & Review*, 26(2), 452–467. <https://doi.org/10.3758/s13423-018-1558-y>
- Scheil, J., & Liefoghe, B. (2018). Motor command inhibition and the representation of response mode during motor imagery. *Acta Psychologica*, 186, 54–62. <https://doi.org/10.1016/j.actpsy.2018.04.008>
- Schieppati, M., Trompetto, C., & Abbruzzese, G. (1996). Selective facilitation of responses to cortical stimulation of proximal and distal arm muscles by precision tasks in man. *The Journal of Physiology*, 491(2), 551–562. <https://doi.org/10.1113/jphysiol.1996.sp021239>
- Schönbrodt, F. D., & Wagenmakers, E.-J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review*, 25(1), 128–142. <https://doi.org/10.3758/s13423-017-1230-y>
- Schönbrodt, F. D., Wagenmakers, E.-J., Zehetleitner, M., & Perugini, M. (2017). Sequential hypothesis testing with Bayes factors: Efficiently testing mean differences. *Psychological Methods*, 22(2), 322–339. <https://doi.org/10.1037/met0000061>
- Silva, A. B., Liu, J. R., Zhao, L., Levy, D. F., Scott, T. L., & Chang, E. F. (2022). A neurosurgical functional dissection of the middle precentral gyrus during speech production. *The Journal of Neuroscience*, 42(45), 8416–8426. <https://doi.org/10.1523/jneurosci.1614-22.2022>
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). A 21 word solution. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2160588>
- Sokolov, A. (1972). *Inner speech and thought*. Springer-Verlag.
- Spieser, L., Aubert, S., & Bonnard, M. (2013). Involvement of SMAp in the intention-related long latency stretch reflex modulation: A TMS study. *Neuroscience*, 246, 329–341. <https://doi.org/10.1016/j.neuroscience.2013.05.005>
- Tian, X., Ding, N., Teng, X., Bai, F., & Poeppel, D. (2018). Imagined speech influences perceived loudness of sound. *Nature Human Behaviour*, 2(3), 225–234. <https://doi.org/10.1038/s41562-018-0305-8>
- Tian, X., & Poeppel, D. (2012). Mental imagery of speech: Linking motor and perceptual systems through internal simulation and estimation. *Frontiers in Human Neuroscience*, 6. <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., Zarate, J. M., & Poeppel, D. (2016). Mental imagery of speech implicates two mechanisms of perceptual reactivation. *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior*, 77, 1–12. <https://doi.org/10.1016/j.cortex.2016.01.002>
- Van Geert, E., Moors, P., Haaf, J., & Wagemans, J. (2022). Same stimulus, same temporal context, different percept? Individual differences in hysteresis and adaptation when perceiving multistable dot lattices. *I-Perception*, 13(4), Article 204166952211093. <https://doi.org/10.1177/20416695221109300>
- Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, 14(5), 779–804. <https://doi.org/10.3758/BF03194105>
- Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, 11, 3571–3594.
- Watkins, K. E., & Paus, T. (2004). Modulation of motor excitability during speech perception: The role of broca's area. *Journal of Cognitive Neuroscience*, 16(6), 978–987. <https://doi.org/10.1162/0898929041502616>
- Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, 41(8), 989–994. [https://doi.org/10.1016/S0028-3932\(02\)00316-0](https://doi.org/10.1016/S0028-3932(02)00316-0)
- Werhahn, K. J., Kunesch, E., Noachtar, S., Benecke, R., & Classen, J. (1999). Differential effects on motorcortical inhibition induced by blockade of GABA uptake in humans. *The Journal of Physiology*, 517(2), 591–597. <https://doi.org/10.1111/j.1469-7793.1999.0591t.x>
- Zhao, L., Silva, A. B., Kurteff, G. L., & Chang, E. F. (2023). Inhibitory control of speech production in the human premotor frontal cortex. <https://doi.org/10.1101/2023.03.01.530698>