Journal of Neural Engineering

PAPER

Imagined speech increases the hemodynamic response and functional connectivity of the dorsal motor cortex

To cite this article: Xiaopeng Si *et al* 2021 *J. Neural Eng.* **18** 056048

View the article online for updates and enhancements.

You may also like

- Decoding human mental states by whole-head EEG+fNIRS during category fluency task performance
Ahmet Omurtag, Haleh Aghajani and Hasan Onur Keles
- A hybrid BCI based on EEG and fNIRS signals improves the performance of decoding motor imagery of both force and speed of hand clenching
Xuxian Yin, Baolei Xu, Changhao Jiang et al.
- fNIRS-GANs: data augmentation using generative adversarial networks for classifying motor tasks from functional near-infrared spectroscopy
Tomoyuki Nagasawa, Takanori Sato, Isao Nambu et al.

This content was downloaded from IP address 90.197.63.230 on 17/05/2022 at 18:51

Journal of Neural Engineering

PAPER

Imagined speech increases the hemodynamic response and functional connectivity of the dorsal motor cortex

Xiaopeng Si[1,2,3,4], Sicheng Li[1,2], Shaoxin Xiang[2,3], Jiayue Yu[2,3] and Dong Ming[1,2,*]

[1] Academy of Medical Engineering and Translational Medicine, Tianjin University, Tianjin 300072, People's Republic of China
[2] Tianjin Key Laboratory of Brain Science and Neural Engineering, Tianjin University, Tianjin 300072, People's Republic of China
[3] Tianjin International Engineering Institute, Tianjin University, Tianjin 300072, People's Republic of China
[4] Institute of Applied Psychology, Tianjin University, Tianjin 300350, People's Republic of China
[*] Author to whom any correspondence should be addressed.

E-mail: richardming@tju.edu.cn

RECEIVED
17 May 2021

REVISED
6 September 2021

ACCEPTED FOR PUBLICATION
10 September 2021

PUBLISHED
7 October 2021

Keywords: functional near-infrared spectroscopy (fNIRS), covert speech, overt speech, sensorimotor cortex, secondary motor cortex, Broca's area, brain–computer interface (BCI)

Supplementary material for this article is available online

Abstract

Objective. Decoding imagined speech from brain signals could provide a more natural, user-friendly way for developing the next generation of the brain–computer interface (BCI). With the advantages of non-invasive, portable, relatively high spatial resolution and insensitivity to motion artifacts, the functional near-infrared spectroscopy (fNIRS) shows great potential for developing the non-invasive speech BCI. However, there is a lack of fNIRS evidence in uncovering the neural mechanism of imagined speech. Our goal is to investigate the specific brain regions and the corresponding cortico-cortical functional connectivity features during imagined speech with fNIRS. *Approach*. fNIRS signals were recorded from 13 subjects' bilateral motor and prefrontal cortex during overtly and covertly repeating words. Cortical activation was determined through the mean oxygen–hemoglobin concentration changes, and functional connectivity was calculated by Pearson's correlation coefficient. *Main results*. (a) The bilateral dorsal motor cortex was significantly activated during the covert speech, whereas the bilateral ventral motor cortex was significantly activated during the overt speech. (b) As a subregion of the motor cortex, sensorimotor cortex (SMC) showed a dominant dorsal response to covert speech condition, whereas a dominant ventral response to overt speech condition. (c) Broca's area was deactivated during the covert speech but activated during the overt speech. (d) Compared to overt speech, dorsal SMC(dSMC)-related functional connections were enhanced during the covert speech. *Significance*. We provide fNIRS evidence for the involvement of dSMC in speech imagery. dSMC is the speech imagery network's key hub and is probably involved in the sensorimotor information processing during the covert speech. This study could inspire the BCI community to focus on the potential contribution of dSMC during speech imagery.

1. Introduction

The speech brain–computer interface (BCI) could decode the user's verbal intent during speech imagery (also known as covert speech) [1]. Decoding imagined speech from the brain could provide a more natural, user-friendly way for developing the next generation BCI [2], which is becoming a research focus. Considering that the neural features during speech imagery remain unclear, there is an urgent need to explore non-invasive neuroimaging modalities for revealing the neural mechanism and features during speech imagery, thus facilitating the development of non-invasive speech BCI [1].

By using the near-infrared lights to measure hemodynamic changes due to the cortical event-related neural activity [3], the functional near-infrared spectroscopy (fNIRS) provides a new brain imaging method to develop the non-invasive speech BCI [1]. Since fNIRS is not affected by the volume

© 2021 IOP Publishing Ltd

conduction effect of the skull [4], it has a higher spatial resolution compared to scalp electroencephalography (EEG) [5], which makes it suitable for revealing cortical activities during speech processing [4]. In addition, fNIRS is less sensitive to motion artifacts than EEG [6], so it befits the study of speech processing mechanisms in the presence of motor movements during vocalization. Moreover, fNIRS has been successfully applied to neurolinguistic researches in speech perception [7], speech production [8, 9], and speech comprehension [10, 11], and it has also been used for speech BCIs [2, 12, 13]. However, current studies of speech imagery mechanisms based on fNIRS are insufficient. The hemodynamic response and functional connectivity patterns during speech imagery need to be further investigated with fNIRS, which could benefit the non-invasive fNIRS speech BCI.

Current studies propose that speech is processed by a cooperated cortical network with multiple brain areas as nodes [14, 15]. Speech imagery was found to involve multiple brain areas by previous functional magnetic resonance imaging (fMRI) and intracranial EEG (iEEG) studies, including Broca's area, sensorimotor cortex (SMC), secondary motor cortex (M2), and temporal auditory area [16–22]. Although previous fNIRS studies validated that the motor cortex [23] was activated during speech imagery, how Broca's area and the subregions of the motor cortex, i.e. dorsal SMC (dSMC), ventral SMC (vSMC), dorsal M2 (dM2), and ventral M2 (vM2) are responded during speech imagery is still unclear. Moreover, previous fNIRS studies mainly focused on the location of cortical activation and failed to explore the functional network during speech imagery. How these areas are cooperatively worked during speech imagery also remains unclear.

Previous fMRI and iEEG studies found that dM2 and dSMC showed enhanced neural responses during speech and singing imagery [24–26]. Considering that dM2 is related to auditory and motor imagery [27, 28] and verbal working memory [29], and dSMC is responsible for speech production processing [30, 31], we hypothesize that fNIRS could also observe the involvement of dSMC and dM2 during speech imagery. In addition, previous fMRI and iEEG studies have found the involvement of Broca's area during speech imagery [16, 19], and Broca's area is associated with the inhibitory control of speech [32] and motor simulation of speech vocalization [25]. Therefore, we hypothesize that fNIRS could also reveal the involvement of Broca's area during speech imagery.

Moreover, fNIRS could also reveal the functional connectivity for the speech network [15, 33], which could provide cortico-cortical functional connectivity features for speech decoding BCI. For example, a previous fNIRS study found that the functional connectivity within non-motor areas was increased during silent reading compared to reading aloud [9]. In contrast, connectivity between the motor and non-motor cortex was decreased. However, how cortical areas cooperatively work during speech imagery remains largely unclear. Considering that previous fMRI studies have found functional connections between the dorsal and ventral motor cortex during speech imagery [16], we further hypothesize that fNIRS could also observe functional connections among the subregions of the motor cortex during speech imagery.

In this study, to investigate the specific brain regions and the corresponding cortico-cortical functional connectivity features during speech imagery with fNIRS, fNIRS signals were recorded from the bilateral motor cortex and Broca's area of 13 subjects during overtly and covertly repeating words. For revealing how the hemodynamic responses are responded during the overt and covert speech, the oxygen–hemoglobin (Hb) concentration changes from the subregions of the motor cortex (dSMC, vSMC, dM2, and vM2) and Broca's area were calculated and statistically analyzed. For investigating how different speech-related cortical areas work cooperatively during speech imagery, the functional connectivity was calculated by Pearson's correlation coefficient. We speculate that the findings could shed light on the channel layout and feature extractions for developing the next generation of speech decoding BCI with fNIRS.

2. Methods

2.1. Subject information

For the fNIRS experiment, 13 healthy participants (ages 22.6 ± 1.4 years old, ages 23–27 years old, ten males and three females) were recruited. All subjects were native Chinese speakers. All subjects had normal intelligence and no history of mental illness. After confirming that each subject met all the inclusion criteria, all participants signed written informed consent. The experiment was reviewed and approved by the Institutional Review Board and Ethics Committee of Tianjin University.

2.2. Experimental paradigm

2.2.1. Experiment preparation

Before the experiment, subjects were seated on a comfortable armchair with 1 m distance in front of a 27 inch LED screen. The experimental paradigm (figure 1) was programmed by the MATLAB-based PsychToolbox [34]. In the beginning, subjects were asked to relax and look at the central fixation cross on the screen for 30 s. Then, subjects were instructed to perform a word production task, alternating overt speech and covert speech trials. At the beginning of the overt/covert trial, a text cue with a 2 s duration was presented on the central screen. The entire experiment included 40 trials with 20 overt trials and

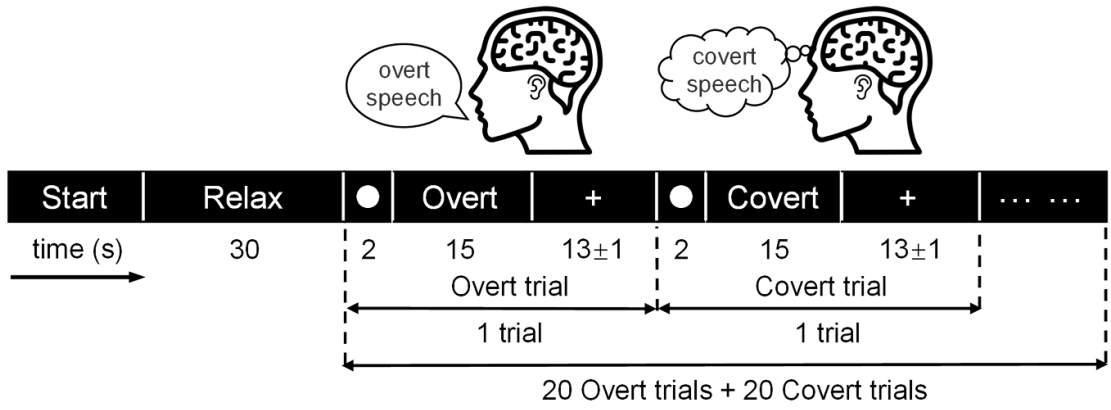

Figure 1. The fNIRS experimental paradigm. There was a 30 s rest before the overt and covert speech trials, and the structure of overt trials and covert trials were equivalent. During the first 2 s, either a 'Ready to start overt speech' or a 'Ready to start covert speech' text was presented to the subject for the following task. Subsequently, the subject was required to overtly or covertly repeat a two-syllable word for 15 s. At last, the subject was asked to keep relaxed and look at the crossing on the screen for 13 s with a jitter of ±1 s.

20 covert trials. The experiment was divided into two runs, each consisting of ten overt and ten covert trials. During each run, subjects were asked to avoid moving their heads and bodies. Between the two runs, subjects were allowed to move and take a short break.

2.2.2. Overt trial

The overt trial included a 2 s task preparation phase, a 15 s overt speech phase, and a rest phase (13 s with ±1 s jitter). Firstly, at the preparation phase, the text cue 'Ready to start overt speech' in Chinese was presented on the screen for 2 s. Subsequently, when the text 'Start overt speech' was shown on the screen, subjects were instructed to repeat a two-syllable Chinese word during the overt speech phase. The duration of the overt speech task was set with 15 s to induce a stronger speech-related hemodynamic response [35]. During this stage, all subjects were asked to repeat the word with normal speech speed and volume in order to keep breathing evenly and avoid brief hypoxia due to speech [36]. Finally, subjects were asked to relax and look at the central fixation cross on the screen during the rest period (13 s with ±1 s jitter).

2.2.3. Covert trial

The structure of the covert trial was the same as that of the overt trial, which consisted of a 2 s task preparation phase, a 15 s covert speech phase, and a rest phase (13 s with ±1 s jitter). During the preparation phase, the text 'Ready to start covert speech' in Chinese was presented on the screen for 2 s. During the covert speech phase, subjects were instructed to imagine speaking the same two-syllable Chinese word actively and repeatedly. The subjects were instructed to imagine moving their articulators and feeling the movement, similar to overt speech without articulation [37, 38]. During the rest phase, subjects were asked to relax and look at the central fixation cross on the screen. The audible vocalization and articulator movements were controlled and reduced as much as possible in three ways. First, subjects were asked to suppress their articulatory movements beforehand. Second, subjects were required to train both overt and covert speech beforehand to suppress muscle movements as much as possible. Third, the subjects' execution of the covert speech was checked during the experiment. We listened to the subjects' voice and occasionally checked whether they had obvious orofacial movements by observing their mouth and neck from the side-view. During our experiments, no subjects were found to have apparent articulatory motor behavior.

2.3. fNIRS data acquisition

2.3.1. fNIRS equipment

The fNIRS data were recorded with a multi-channel fNIRS system (NirScan, Danyang Huichuang Medical Equipment Co. Ltd) with a sampling rate of 9 Hz (figure 2(a)). The photo of the experiment was used with the consent of this subject. Near-infrared light of three different wavelengths (850, 808, and 740 nm) was used to measure the hemodynamic responses of the Hb concentration, including the oxygenated Hb (HbO) and deoxygenated Hb (HbR) concentration. For the fNIRS channel configuration, 15 light sources and 16 detectors were plugged into a holder on the standard EEG cap by referring to the standard international 10–20 system of electrode placement, forming 48 measurement channels (figure 2(b)). The distance between a light source and detector was 3 cm.

2.3.2. fNIRS channel location and cortical sensitivity

To obtain each fNIRS channel's Montreal Neurological Institute (MNI) coordinate, the spatial coordinates of light sources, detectors, and anchor points (located at Nz, Cz, Al, Ar, Iz in the 10–20 system of electrode placement) were firstly measured

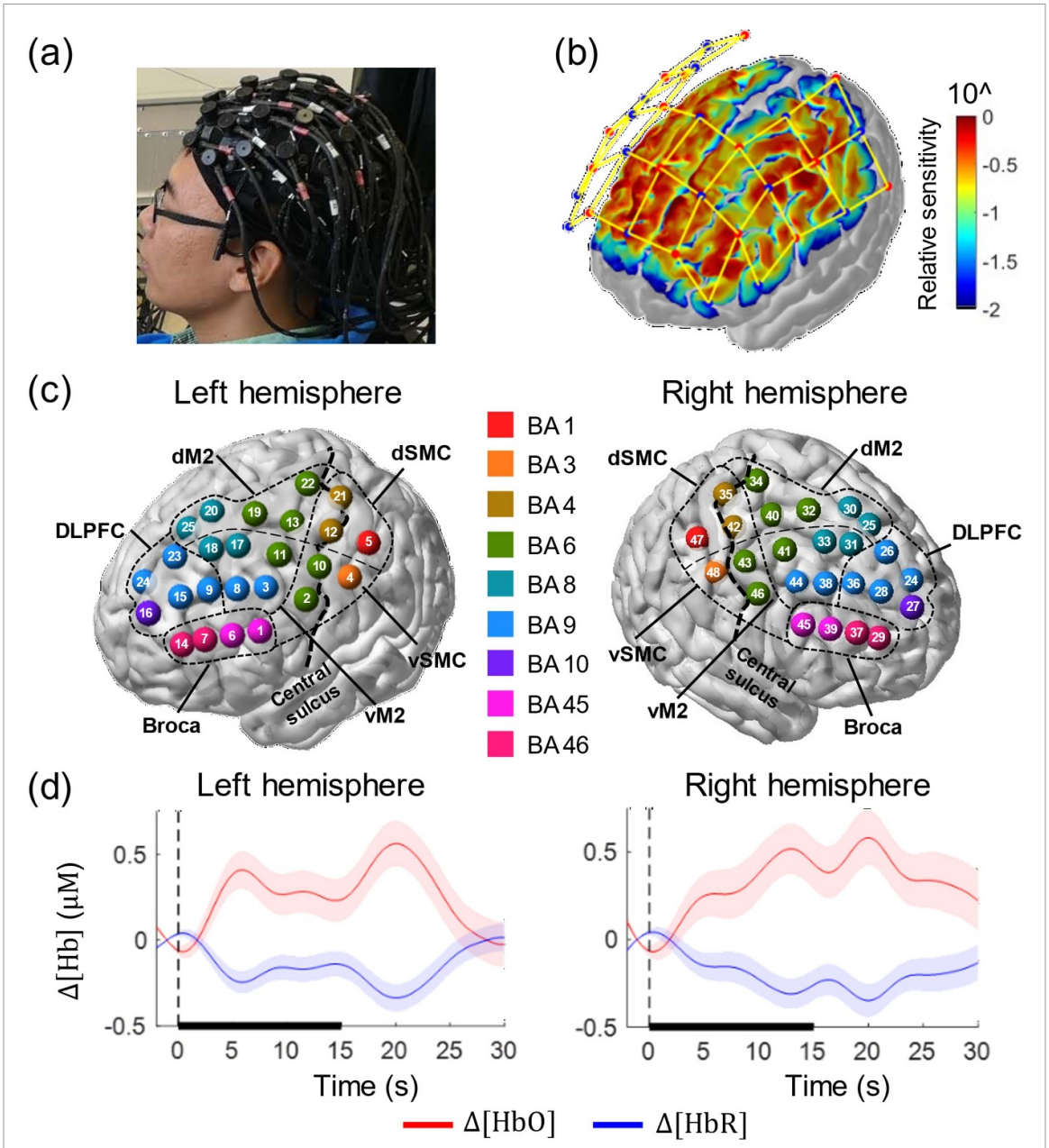

Figure 2. The fNIRS measurements and grand-averaged hemodynamic responses. (a) A photo of a subject wearing an fNIRS measuring cap with 48 channels. (b) fNIRS channel configuration. The red points indicated the light sources, the blue points indicated the detectors, and the yellow lines indicated the fNIRS channels. The color on the cerebral cortex indicated logarithmic relative sensitivity of fNIRS to cortical hemodynamic response, ranging from 0.01 to 1. (c) Anatomical location. The location of the channels and the corresponding Brodmann labels were shown on the Colin27 standard brain template. The circles represented the location where the center of the channel is projected to the cortex, and the color represented the corresponding Brodmann (BA) areas. The black dotted lines circled 6 (ROIs): Broca's area (Broca), ventral and dorsal sensorimotor cortex (vSMC and dSMC), ventral and dorsal secondary motor cortex (vM2 and dM2), as well as the DLPFC. The thick black dotted line represented the central sulcus. (d) Grand-averaged Δ [HbO] and Δ [HbR] responses across all subjects were shown for typical channels (channel 2 from left vSMC and channel 46 from right vSMC).

by using the electromagnetic 3D digitizer system (FASTRAK. Polhemus VT, USA). Secondly, the spatial registration approach registered the measured spatial coordinates to the standard MNI space [39]. Thirdly, the MNI coordinate of the midpoint between a light source and detector pair was determined, as each fNIRS channel consisted of a light source and detector pair. Fourthly, the midpoint was spatially projected to the cortex surface of the 'Colin27' brain template [40], and the fNIRS channel's coordinate was defined as the MNI coordinate of the cortical projected point [39]. Finally, to reduce the measurement error of fNIRS channel coordinates, the *x*, *y* and *z* values of the left and right hemispheric symmetric channels were averaged and marked with positive or negative hemispheric information (table 1).

To determine the cortical area to which each fNIRS channel was theoretically sensitive, the Monte–Carlo photon transmitted software tMCimg [41] was used to simulate the probability distribution of

Table 1. MNI coordinates and Brodmann labels of channel cortical projection points.

Channel	*X*	*Y*	*Z*	Brodmann label	Region of interest
Ch1/Ch45	−54/54	31	27	45	Broca
Ch2/Ch46	−59/59	8	39	6	vSMC
Ch3/Ch44	−46/46	22	44	9	vM2
Ch4/Ch48	−60/60	−16	44	3	vSMC
Ch5/Ch47	−54/54	−31	55	1	dSMC
Ch6/Ch39	−47/47	42	25	45	Broca
Ch7/Ch37	−43/43	55	25	46	Broca
Ch8/Ch38	−41/41	35	45	9	vM2
Ch9/Ch36	−31/31	44	43	9	DLPFC
Ch10/Ch43	−54/54	−3	51	6	vSMC
Ch11/Ch41	−44/44	13	58	6	vM2
Ch12/Ch42	−45/45	−16	62	4	dSMC
Ch13/Ch40	−32/32	−3	62	6	dM2
Ch14/Ch29	−32/32	60	19	46	Broca
Ch15/Ch28	−23/23	55	39	9	DLPFC
Ch16/Ch27	−14/14	67	30	10	DLPFC
Ch17/Ch33	−29/29	26	60	8	vM2
Ch18/Ch31	−21/21	36	58	8	vM2
Ch19/Ch32	−24/24	13	71	6	dM2
Ch20/Ch30	−5/5	24	67	8	dM2
Ch21/Ch35	−34/34	−29	71	4	dSMC
Ch22/Ch34	−26/26	−16	79	6	dM2
Ch23/Ch26	−10/10	49	53	9	DLPFC
Ch24 (Middle)	2	50	34	9	DLPFC
Ch25 (Middle)	3	32	54	8	dM2

Broca = Broca's area, including BA 45 and its adjacent BA 46; vSMC = ventral sensorimotor cortex, including precentral and postcentral gyrus; dSMC = dorsal sensorimotor cortex; vM2 = ventral secondary motor cortex, a joint region of premotor cortex (PMC), supplementary motor area (SMA) and frontal eye field (FEF); dM2 = dorsal secondary motor cortex; and DLPFC = dorsolateral prefrontal cortex, including dorsal BA 9 and BA 10.

photon transmission across the cortex for each light source and detector pair. Then, the cortical sensitivity was visualized on the cortical surface of the 'Colin27' brain template by using the AtlasViewer [41] (figure 2(b)).

2.3.3. Brodmann atlas (BA) label and region of interest (ROI)

To obtain the corresponding BA label of each channel, the MNI-to-BA tool [42] was used. Firstly, a 10 mm radius sphere with each fNIRS channel's MNI coordinate as the center was determined. Secondly, all possible BA labels from cortical voxels within the sphere were determined. Then, the percentage of each BA label's voxels over all the cortical voxels within the sphere was calculated. Next, the BA label with the highest percentage was identified as the BA label of a given fNIRS channel. Finally, each fNIRS channel and the corresponding BA label were visualized on the 'Colin27' brain template by using BrainNet Viewer [43].

According to each fNIRS channel's BA label and the anatomical location information from the 'Colin27' brain template, all 48 fNIRS channels were divided into the following six regions of interests (ROIs) (figure 2(c), table 1), including Broca's area (Broca) [44], vSMC [23], dSMC [24], vM2 [45, 46], dM2 [45, 47], and dorsolateral prefrontal cortex (DLPFC) [48]. Broca was located in the inferior frontal gyrus, including BA 45 and its adjacent BA 46 [44]. SMC included the precentral and postcentral gyrus (BA 1, BA 3, BA 4, and BA 6) [23]. M2 included the premotor cortex, the supplementary motor cortex, and the frontal eye field (BA 6, BA 8, and BA 9) [45]. DLPFC was located in the superior frontal gyrus and middle frontal gyrus (BA 9 and BA 10) [48]. The regional boundaries (in figure 2(c)) were drawn according to the detectable cortical range of fNIRS (i.e. sensitivity map, figure 2(b)).

2.4. fNIRS data pre-processing

To remove the systematic, non-event related noise components [49] and obtain the high signal-to-noise ratio (SNR) HbO and HbR signals, the raw fNIRS data were pre-processed by the following steps. (a) Channels with poor signal quality were excluded from further analysis if the raw data met either one of the following criteria. 1. The amplitude of the optical intensity (fNIRS raw data) exceeded the range of 0.5–1000; 2. The SNR of optical intensity was smaller than the threshold of 2 [50], in which the SNR was defined as the standard deviation over the mean value for optical intensity. If more than 25% of channels of a subject were excluded, then the subject would be excluded from further analysis (table 2). This is a reliable threshold [51–53] to exclude subjects

Table 2. Subject information.

Subject	Gender	Age (years)	Excluded channels/all	Subject accepted (≤25% channels excluded)
1	Female	23	2/48	Yes
2	Male	22	12/48	Yes
3	Male	22	12/48	Yes
4	Male	22	2/48	Yes
5	Male	22	9/48	Yes
6	Male	22	8/48	Yes
7	Male	22	4/48	Yes
8	Female	22	6/48	Yes
9	Female	22	7/48	Yes
10	Male	23	4/48	Yes
11	Male	27	8/48	Yes
12	Male	22	45/48	No
13	Male	23	29/48	No

affected by systematic interference such as hair disturbance [54] or individual differences in skull thickness [55] as much as possible. (b) The optical intensity was converted into the optical density signals [56]. (c) To remove the non-neural related response signals such as heart rate (∼1 Hz) and breathing rate (∼0.3 Hz) [57], a three-order 0.01–0.2 Hz bandpass infinite impulse response Butterworth filter was used on the optical density signal [57]. (d) The modified Beer–Lambert law was used to convert the filtered optical density signals into HbO and HbR signals [56]. For our three-wavelength fNIRS, the estimated differential path-length factor parameter was set to six for each wavelength, following the default setting of HomER2 [58]. The above four steps were performed by the MATLAB-based HomER 2 toolkit [56]. (e) To remove the noise components such as Mayer waves (∼0.1 Hz) which cannot be removed by bandpass filtering, the hemodynamic signal separation algorithm was used [59]. By decomposing HbO and HbR data into a negative linear relationship component (functional brain activity signal) and a positive linear relationship component (systemic noise), the algorithm could effectively extract the functional and neural related HbO and HbR signals from the systematic and non-event related physical activity noise [60]. It has been proved that this algorithm could improve HbO and HbR signal quality in previous fNIRS experiments [61–63]. The HbO and HbR signals' functional components were finally extracted by the above algorithm for further data analysis.

2.5. Event-related Hb concentration changes calculation

Event-related Hb concentration changes (Δ[HbO] and Δ[HbR]) for overt/covert conditions were calculated by the following steps. (a) Each trial's event-related Hb concentration changes were baseline corrected by subtracting the mean value of pre-stimulus 2 s duration. (b) After baseline correction, event-related Hb concentration change signals were plotted for different conditions. Firstly, to reveal the fNIRS data quality, event-related Hb concentration changes of grand-averaged data of all subjects across all conditions (overt and covert) were plotted for typical channels from left and right vSMC (figure 2(d)). Then, to reveal different cortical areas' responses in different conditions, event-related Hb concentration changes for the overt and covert conditions were plotted for typical channels, such as left vSMC (figures 3(a)–(c)), right dM2 (figures 3(d)–(f)), and right Broca (figures 3(g)–(i)). To visualize the location of each channel and its effective detection range, the sensitivity distributions of the typical channels were also plotted. In addition, all channels' event-related Hb concentration changes were plotted for overt (figure S1 (available online at stacks.iop.org/JNE/18/056048/mmedia)) and covert (figure S2) conditions, respectively.

2.6. The statistical mapping of mean HbO changes between speech tasks and baseline conditions

Since HbO and HbR are negatively correlated, the information of HbR signals can also be reflected by HbO signals [59, 62]. Therefore, only HbO signals were used in the following sections. To compare the strength of event-related HbO concentration changes between tasks and baseline conditions, the following steps were performed. (a) The post-stimulus 25 s HbO concentration changes were averaged as the mean HbO changes under overt or covert conditions for each trial. (b) The pre-stimulus 2 s HbO concentration data were averaged as the mean HbO changes for each trial under baseline conditions. (c) To compare the strength of event-related HbO concentration changes between tasks (covert and overt) and baseline conditions, the paired tests were conducted according to Wilcoxon signed-rank test on mean HbO changes data for each channel. This nonparametric test does not require the residuals of the fitted model to follow a normal distribution (i.e. the within-group variation of the sample group should be assumed to follow a normal distribution) and is therefore more flexible [64]. (d) To minimize the risk

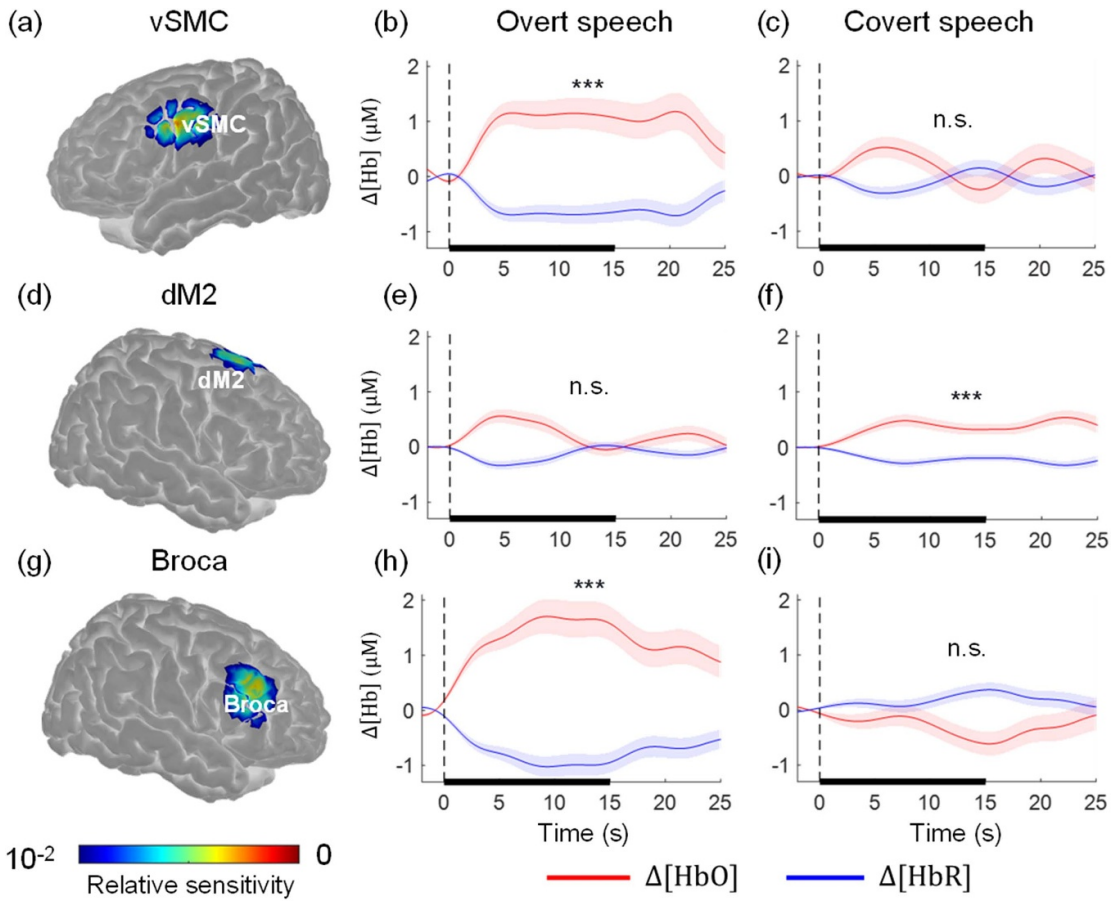

Figure 3. The typical cortical response which was different during the overt and covert speech. (a) Sensitivity distribution of a typical channel in vSMC (channel 2). (b) and (c) The time course in Δ [HbO] and Δ [HbR] of the vSMC channel, where the mean HbO changes showed a significant increase during the overt speech but not during the covert speech (significance was tested only for Δ [HbO] response, Wilcoxon signed-rank test, FDR corrected, *** $p < 0.0005$, ** $p < 0.005$, * $p < 0.05$, n.s. non-significant, the same below). (d) Sensitivity distribution of a typical channel in dM2 (channel 30). (e) and (f) The time course in Δ [HbO] and Δ [HbR] of the dM2 channel, where the mean HbO changes showed a significant increase during the covert speech but not during the overt speech. (g) Sensitivity distribution of a typical channel in Broca (channel 39). (h) and (i) The time course in Δ [HbO] and Δ [HbR] of the Broca channel, where the mean HbO changes showed a significant increase during the overt speech and a decreasing trend during the covert speech.

of type I errors, the *p*-values were corrected with false discovery rate (FDR) correction for multiple comparisons and were labeled on the time course of HbO concentration changes signals (figure 3). Moreover, to reduce the risk of type II errors, the effect size and the statistical power were post-hoc tested using G*Power. Previous studies have found that statistical power below 0.8 are acceptable for small and medium effect sizes. For example, Szucs *et al* found that for small effect size (Cohen's $d < 0.2$), the median statistical power of cognitive neuroscience studies was 0.11; for medium effect size ($0.2 < d < 0.5$), the median statistical power of cognitive neuroscience studies was 0.40 [65]. In addition, two studies assumed that type II error $\beta = 0.8$ (i.e. power = 0.2) was sufficient for small effect size [66, 67]. In this study, if a channel's statistical power was less than 0.25 with a small effect size [65–67], or less than 0.40 with a medium effect size [65], the corresponding statistical result was postulated to have insufficient power and excluded. (e) Finally, to reveal the cortical responses during speech tasks (overt or covert), the mean HbO changes (figures 4(a) and 5(a)) and FDR corrected *p*-values (figures 4(b) and 5(b)) between task and baseline conditions were visualized on the Colin27 [40] standard brain template by AtlasViewer [41] for each channel. (f) In addition, the significant channels ($p < 0.05$) for overt speech (figure S3(a)) and covert speech (figure S3(b)) conditions were marked on the brain template.

2.7. Comparison of region-wise HbO responses

We did the following analysis to quantify the response differences in the five speech-related cortical regions (dSMC, vSMC, dM2, vM2, Broca) between the two speech conditions (covert and overt speech). First, region-wise cortical response patterns were statistically compared between covert and overt speech. The null hypothesis of the statistics was that the neural processes in each region were the same for covert and overt speech. The steps of the statistical analysis were as follows. (a) The region's speech-related channels

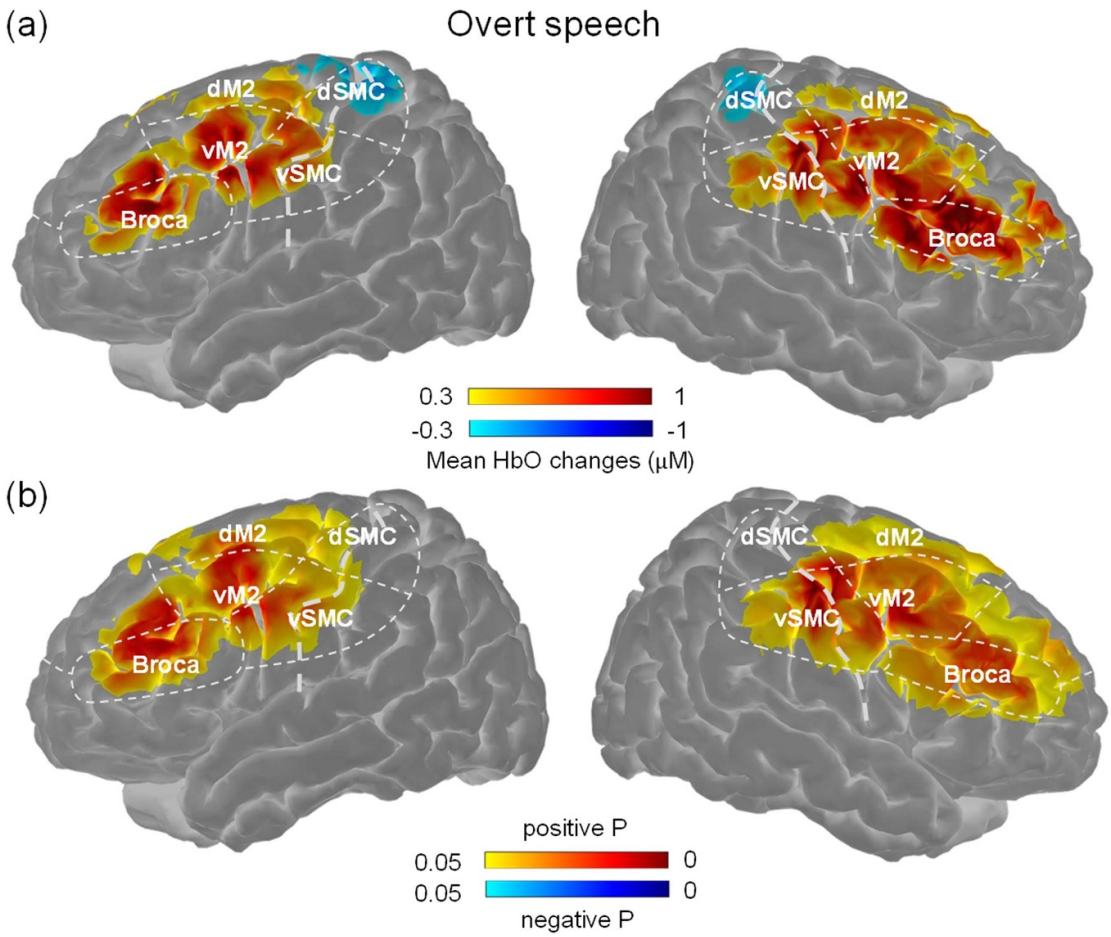

Figure 4. The mean HbO concentration changes and statistical mapping for overt speech condition. (a) Mean HbO changes during the overt speech of each channel. (b) Significance of mean HbO changes during the overt speech ($p < 0.05$, Wilcoxon signed-rank test, FDR corrected, $n = 48$). A warm color indicated the response was higher than the baseline, while a cold color indicated the opposite way. The dotted line represented the area boundary.

were defined as the union of covert and overt significant channels. DLPFC was excluded from the comparison due to the absence of significantly activated channels during imagery speech there. (b) The mean HbO changes in a region during the overt speech were calculated as the average of all speech-related channels' mean HbO changes, and the mean HbO changes during the covert speech were calculated in the same way. (c) To compare the region-wise response patterns, the two-sample tests were conducted according to Wilcoxon rank-sum test on the mean HbO changes during the overt vs. covert speech in each region, and the *p*-values were corrected with FDR correction for multiple comparisons (figures 6(a)–(c)). (d) The same calculation method was also used to calculate the interhemispheric differences in region-wise responses (figure S4).

Second, the extent of engagement between the dorsal and ventral area was compared by calculating speech-related areas' dorsal and ventral response percentages. The dorsal/ventral response percentage was defined as the number of significant channels in a region's dorsal/ventral part, divided by the total number of significant channels within the region. For example, the ventral response percentage of M2 for overt speech condition was calculated as the proportion of significant channels in bilateral vM2 (figure S3(a), seven channels) to all significant channels in M2 (figure S3(a), 11 channels) during the overt speech. A higher percentage (>50%) means a more wide-ranging engagement compared to its counterpart during the overt or covert speech (figures 6(d) and (e)).

2.8. Statistical mapping of mean HbO data between covert and overt speech conditions

To compare the strength of event-related HbO concentration changes between covert and overt speech conditions, the following steps were performed. (a) The mean value of post-stimulus 25 s HbO concentration data was calculated as the mean HbO changes for each trial under overt or covert speech conditions. (b) The difference value of mean HbO changes between covert and overt conditions was visualized on the Colin27 standard brain template by AtlasViewer for each channel (figure 7(a)). (c) To compare the difference, the non-parameter rank-sum tests were conducted on mean HbO changes data

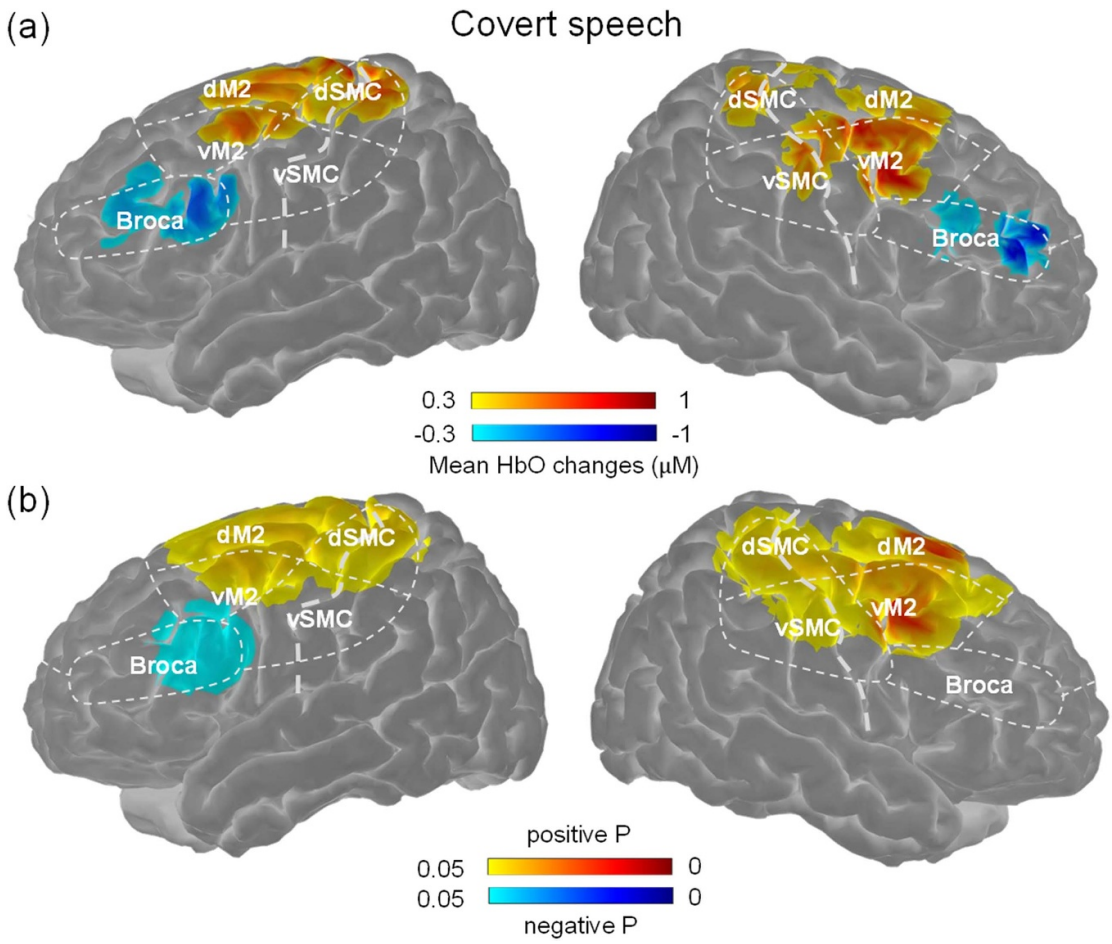

Figure 5. The mean HbO concentration changes and statistical mapping for covert speech condition. (a) Mean HbO changes during the covert speech of each channel. (b) Significance of mean HbO changes during the covert speech ($p < 0.05$, Wilcoxon signed-rank test, FDR corrected, $n = 48$). A warm color indicated the response was higher than the baseline, while a cold color indicated the opposite way. The dotted line represented the area boundary.

between covert and overt speech conditions. To minimize the risk of type I errors, the p-values were corrected with Bonferroni correction for multiple comparisons. They were visualized on the brain template for each channel (figure 7(b)). In addition, the significant channels with $p < 0.05$ were marked on the brain template (figure S5).

2.9. Functional connectivity and statistical analysis

Speech is considered to be processed with distributed cooperative networks [14, 15]. After calculating the network's nodes activities (i.e. cortical regions) differences between overt and covert speech, we wanted to further explore the differences in the network's edges (i.e. the interactions between cortical regions). To reflect the synchronization of cortical activity between regions [68], functional connectivity between speech-related regions (vSMC, dSMC, vM2, dM2, Broca) during the overt and covert speech was calculated and statistically tested. (a) To obtain pairwise channels' functional connectivity during the overt speech for each subject, the HbO concentration time series data of the significant speech responsive channels (i.e. the union of the overt and covert significant channels) were paired off to calculate the Pearson's correlation coefficient. This metric has been considered in several previous studies to be reliable for quantifying the resting-state [69] and task-state [70] functional networks. The HbO concentration time series were the post-stimulus 25 s concentration changes, the same as the calculation of activation in section 2.6. (b) To further determine channel-to-channel significant functional connectivity before calculating region-to-region functional connectivity, nonparametric permutation tests were used referring to other studies of functional networks [15]. A 500-times permutation test (the channel pairs' corresponding trials were shuffled randomly) was performed, and the significant functional connectivity was extracted through the threshold $p < 0.05$. Significant functional connectivity indicated the reliability for reflecting speech information processing rather than interference. (c) To obtain the group level's region-to-region functional connectivity, the mean r-values of significant channel-to-channel correlations were defined as the connectivity's strength between ROIs (figure S6(a)), which were finally visualized on the standard brain template

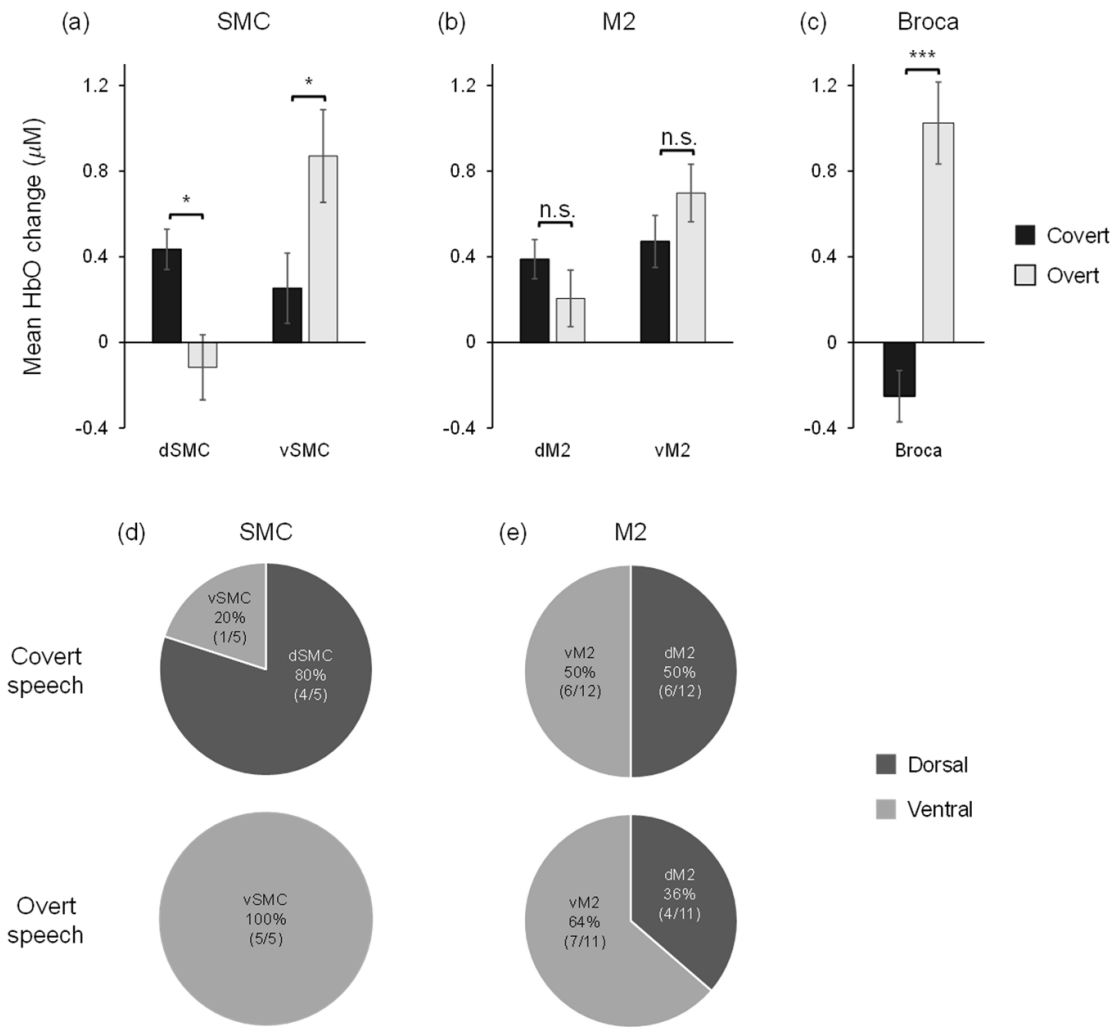

Figure 6. Comparison of region-wise cortical response patterns. (a) The mean HbO changes during the covert speech were significantly stronger than during the overt speech in the dSMC (mean ± sem across channels × subjects, $p < 0.05$, Wilcoxon rank-sum test, FDR corrected, the same below). In contrast, the mean HbO changes during the overt speech were significantly stronger in the vSMC than during the covert speech. (b) M2 showed a similar trend as SMC without significance. (c) In Broca, mean HbO changes during the overt speech were stronger than those during the covert speech. (d) SMC showed a larger extent of dorsal engagement to covert speech condition, whereas a larger extent of ventral engagement to overt speech condition. (e) M2 also showed a larger extent of ventral engagement to overt speech condition, but the ventral ratio was smaller than SMC. The ratio of ventral or dorsal significant channels to all significant channels within a region was indicated in parentheses.

(figure 8(a)). (d) Similarly, the group level's region-to-region functional connectivity during the covert speech was calculated, statistically tested (figure S6(b)), and visualized following the steps (a)–(c) (figure 8(b)). (e) To investigate if there exists a significant connectivity strength difference between covert and overt speech, Wilcoxon rank-sum tests were conducted for all region-to-region connections, and FDR corrected. The connections with significant contrast were colored (figure 8(b)).

3. Results

3.1. Grand-averaged HbO and HbR concentration changes

The grand-averaged HbO and HbR concentration change signals across all subjects on the typical channels from the bilateral vSMC (channel 2 from left hemisphere and channel 46 from right hemisphere) showed that: (a) The HbO concentration change signals had a slow increase initially, reached a plateau after 5 s, and started to decrease about 5 s after the end of the task. (b) The HbR concentration change signals had an opposite trend to the HbO concentration change signals (figure 2(d)).

3.2. Typical channels which showed significant hemodynamic response during the overt and covert speech

Significant activations were shown from typical channels in motor cortex and Broca's area during the overt and covert speech. Firstly, for the typical channel from vSMC (figure 3(a), channel 2), the overt condition's task-related HbO concentration changes signal were consistently higher than baseline, and the mean HbO changes were significantly higher than baseline

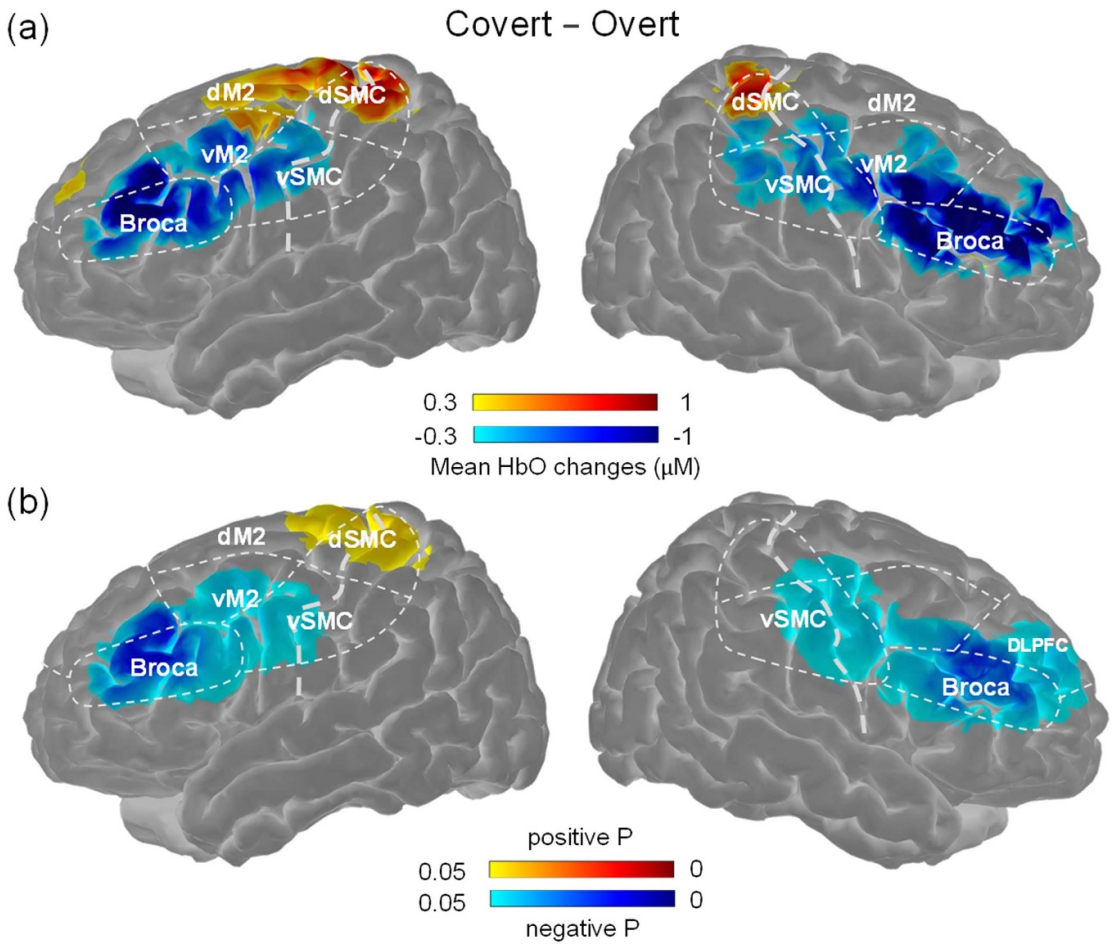

Figure 7. The mean HbO concentration changes and statistical mapping for the contrast between covert and overt speech conditions. (a) Mean HbO changes during the overt speech minus overt speech of each channel. (b) The significant differences of mean HbO changes between covert and overt speech conditions ($p < 0.05$, Wilcoxon signed-rank test, FDR corrected, $n = 48$). A warm color indicated the response was higher than the baseline, while a cold color indicated the opposite way. The dotted line represented the area boundary.

($p < 0.0005$, Wilcoxon signed-rank test, FDR corrected, the same below; mean = 0.94 μM; figure 3(b)). The covert condition's task-related HbO concentration changes did not show significant changes (n.s.; mean = 0.17 μM; figure 3(c)). Secondly, for the typical channel from dM2 (figure 3(d), channel 30), the covert condition's task-related HbO concentration changes were consistently higher than baseline, and the mean HbO changes were significantly higher than baseline ($p < 0.0005$; mean = 0.36 μM; figure 3(f)). The overt condition's task-related HbO concentration changes did not show significant changes (n.s.; mean = 0.22 μM; figure 3(e)). Thirdly, for the typical channel from Broca (figure 3(g), channel 39), the overt condition's task-related HbO concentration changes were consistently higher than baseline, and the mean HbO changes were significantly higher than baseline ($p < 0.0005$; mean = 1.28 μM; figure 3(h)). The covert condition's task-related HbO concentration changes were consistently lower than baseline without significance (n.s.; mean = −0.30 μM; figure 3(i)). In summary, the above hemodynamic response results suggested that overt speech and covert speech showed different cortical response patterns. The response trend of each region would be further analyzed in the following sections.

3.3. The bilateral vSMC, vM2 and Broca were significantly activated during the overt speech

To further explore the cortical response patterns in the prefrontal and motor areas during the overt speech, the mean HbO concentration changes (figure 4(a)) and statistical mapping (figure 4(b)) results were shown. Firstly, there were increased mean HbO changes for bilateral vSMC and vM2 (figure 4(a)), which were also significantly activated (Wilcoxon signed-rank test, FDR corrected, $n = 48$, $p < 0.05$; figure 4(b)). There were 12 significantly activated channels (channel 2, 10, 43, 46, 48 from bilateral vSMC and channel 3, 8, 17, 33, 38, 41, 44 from bilateral vM2; figure S3(a)). Secondly, for the bilateral Broca, there also existed increased mean HbO changes during the overt speech (figure 4(a)), which was also significantly activated (figure 4(b)). There were six significantly activated channels from bilateral Broca (channel 7, 8, 29, 37, 39, 45) and

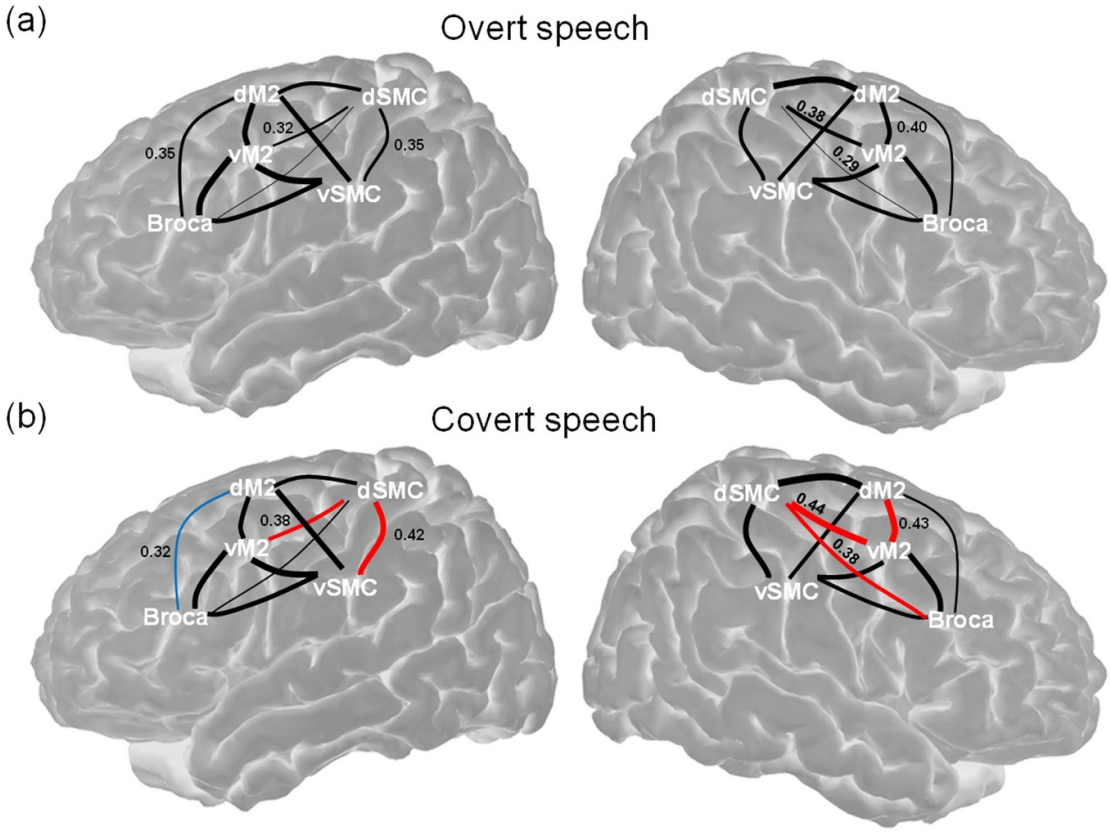

Figure 8. Statistical comparison of the functional connectivity between covert and overt speech conditions. (a) During the overt speech, there were significant inter-regional functional connections between Broca, vM2, dM2, vSMC, and dSMC. (b) The functional connections between dSMC and other regions became stronger, and the functional connection between dM2 and Broca became weaker during the covert speech. The lines indicated the significant Pearson's correlation coefficient between regions ($p < 0.05$, permutation test). The colored lines represented the significantly different functional connections between regions during the covert than overt speech ($p < 0.05$, Wilcoxon rank-sum test, FDR corrected, $n = 20$), with red represented enhanced connections and blue represented weakened connections. The numbers indicated the Pearson's correlation coefficient r-values of the significant connections.

two channels from the right Broca-adjacent region DLPFC (channel 27, 36; figure S3(a)). Thirdly, bilateral dM2 showed increased mean HbO changes during the overt speech (figure 4(a)) with four significantly activated channels (channel 13, 25, 32, 40 from bilateral dM2; figure 4(b)). Whereas bilateral dSMC showed decreased mean HbO changes without significance (figure 4). In summary, the bilateral vSMC, vM2 and Broca were most significantly activated, whereas bilateral dM2 was less significantly activated during the overt speech.

3.4. The bilateral dSMC and dM2 were significantly activated, whereas Broca was deactivated during the covert speech

The mean HbO concentration changes (figure 5(a)) and statistical mapping (figure 5(b)) results during the covert speech showed that the motor cortex, especially the dorsal part (dSMC and dM2), was significantly activated. Firstly, there were increased mean HbO changes for bilateral dSMC and dM2 (figure 5(a)), which were also significantly activated (Wilcoxon signed-rank test, FDR corrected, $n = 48$, $p < 0.05$; figure 5(b)). There were ten significantly activated channels (channel 12, 21, 35, 42 from bilateral dSMC and channel 13, 19, 20, 22, 30, 32 from bilateral dM2; figure S3(b)). Secondly, for the ventral motor cortex, there were increased mean HbO changes for bilateral vM2 and right vSMC during the covert speech (figure 5(a)), which were also significantly activated (figure 5(b)). There were seven significantly activated channels (channel 11, 17, 31, 33, 41, 44 from bilateral vM2, and channel 43 from right vSMC; figure S3(b)). Thirdly, bilateral Broca showed decreased mean HbO changes during the covert speech (figure 5(a)), in which left Broca was significantly deactivated (figure 5(b)) with one significant channel (channel 1 from left Broca; figure S3(b)). In summary, dSMC and dM2 were significantly activated, whereas Broca was deactivated during the covert speech. The cortical response patterns in these regions were different from overt speech.

3.5. dSMC showed a larger engagement to covert speech, whereas vSMC showed a larger engagement to overt speech

Significances of regional differences were found between covert and overt speech (figures 6(a)–(c)).

Firstly, dSMC's mean HbO changes during the covert speech were significantly higher than that during the overt speech (Wilcoxon rank-sum test, $p < 0.05$, FDR corrected. $n = 5$; covert speech: 0.43 ± 0.09 μM, overt speech: -0.12 ± 0.15 μM, mean $\pm$ sem; figure 6(a)). Whereas, vSMC's mean HbO changes during the covert speech were significantly lower than that during the overt speech ($p < 0.05$; covert speech: 0.25 ± 0.16 μM, overt speech: 0.87 ± 0.22 μM; figure 6(a)). Secondly, dM2's mean HbO changes during the covert speech showed a higher trend than that during the overt speech without significance (n.s.; covert speech: 0.39 ± 0.09 μM, overt speech: 0.21 ± 0.13 μM; figure 6(b)). Whereas, vM2's mean HbO changes during the covert speech showed a lower trend than that during the overt speech without significance (n.s.; covert speech: 0.47 ± 0.12 μM, overt speech: 0.70 ± 0.13 μM; figure 6(b)). Thirdly, Broca's mean HbO changes during the covert speech were also significantly lower than that during the overt speech, in which Broca was deactivated during the covert speech and activated during the overt speech ($p < 0.0005$; covert speech: -0.25 ± 0.12 μM; overt speech: 1.03 ± 0.19 μM; figure 6(c)). Overall, in dSMC, the HbO response during the covert speech was significantly stronger than overt speech; in vSMC and Broca, it was significantly stronger during the overt speech. This may imply that functions of the above regions might be different during the covert and overt speech.

Furthermore, the dorsal motor cortex showed a more wide-ranging engagement trend during the covert speech, while the ventral part showed the same trend during the overt speech. Firstly, for SMC, the proportion of significantly responsive channels in dSMC over all SMC (dSMC and vSMC) was 80%, which showed a dorsal dominance trend during the covert speech. And the proportion of significantly responsive channels in vSMC over all SMC was 100%, which showed a ventral dominance trend during the overt speech (figure 6(d)). Secondly, for M2, the proportion of significantly responsive channels in dM2 over all M2 (dM2 and vM2) was 50%. The proportion of significantly responsive channels in vM2 over all M2 was 64%, which showed a ventral dominance trend during the overt speech (figure 6(e)). Thirdly, for the comparison of M2 and SMC, the proportion of significantly responsive channels during a covert speech in dM2 was lower than in dSMC (dM2: 50%, dSMC: 80%), and the proportion of significantly responsive channels during the overt speech in vM2 was lower than in vSMC (vM2: 64%, vSMC: 100%). In summary, compared to M2, SMC showed a larger extent of dorsal engagement to covert speech condition, whereas a larger extent of ventral engagement to overt speech condition.

3.6. Insignificant hemisphere lateralized response was observed in speech-related areas during the covert and overt speech

To quantify the lateralization of the cortical response, the interhemispheric response differences between the five subregions (vSMC, dSMC, vM2, dM2, and Broca) under the two conditions (figure S4) was calculated. The overt speech showed a slightly right-hemisphere HbO response dominance in vSMC, dM2, vM2, and Broca without statistical significance (figure S4(a), $p > 0.05$). The covert speech response showed a slightly left-hemisphere HbO response dominance in dSMC, dM2 and Broca without significance and a slightly right-hemisphere HbO response dominance in vSMC and vM2 without significance (figure S4(b)).

3.7. dSMC and dM2 showed a stronger HbO response during the covert speech, whereas vSMC, vM2 and Broca showed a weaker HbO response compared to overt speech

After region-wise analysis, the differences between covert and overt speech for each channel was further compared. The difference of the HbO responses within each region was presented by mapping the mean HbO concentration changes contrast between covert and overt speech conditions (figure 7(a)) and statistical significance (figure 7(b)). These results revealed that dSMC and dM2 showed significantly increased responses during the covert speech compared to overt speech, whereas vSMC, vM2 and Broca from the bilateral hemisphere showed significantly decreased responses during the covert speech compared to overt speech. Firstly, there were increased mean HbO concentration differences for bilateral dSMC and left dM2 during the covert speech compared to overt speech (figure 7(a)). And the left dSMC and dM2 showed significantly increased response (Wilcoxon signed-rank test, FDR corrected, $n = 48$, $p < 0.05$; figure 7(b)). There were two significantly increased channels (channel 21 from left dSMC, and channel 22 from left dM2; figure S5). Secondly, there were decreased mean HbO concentration differences for bilateral vSMC and left vM2 (figure 7(a)), which also showed significantly decreased responses during the covert speech compared to overt speech (figure 7(b)). There were four significantly decreased channels (channel 2, 43, 46 from bilateral vSMC and channel 3 from left vM2; figure S5). Thirdly, there were also decreased mean HbO concentration differences for bilateral Broca (figure 7(a)), which also showed significantly decreased responses during the covert speech compared to overt speech (figure 7(b)). There were seven significantly decreased channels from bilateral Broca (channel 1, 6, 7, 29, 37, 39, 45) and one channel from the right Broca-adjacent region DLPFC (channel 27; figure S5). Taking together, the

dorsal motor cortex preferred to covert speech condition, whereas the ventral motor cortex and Broca preferred to overt speech condition.

3.8. Compared to overt speech condition, dSMC-related functional connections were enhanced for covert speech condition

The functional connectivity of overt speech (figure 8(a)) and covert speech conditions (figure 8(b)) results showed that the dSMC-related functional connections were enhanced during the covert speech compared to overt speech. Firstly, there were five significantly enhanced functional connections during the covert speech compared to overt speech ($p < 0.05$, FDR corrected, left hemisphere: dSMC-vSMC, $r = 0.423$ during the covert speech, $r = 0.353$ during the overt speech; dSMC-vM2, $r = 0.380$ during the covert speech, $r = 0.320$ during the overt speech; right hemisphere: dSMC-vM2, $r = 0.440$ during the covert speech, $r = 0.383$ during the overt speech; dM2-Broca, $r = 0.382$ during the covert speech, $r = 0.286$ during the overt speech; dM2-vM2, $r = 0.433$ during the covert speech, $r = 0.401$ during the overt speech; figures 8(b) and S6). Besides, there was one significantly diminished functional connection during the covert speech compared to overt speech ($p < 0.05$, FDR corrected, left hemisphere: Broca-dM2, $r = 0.316$ during the covert speech, $r = 0.347$ during the overt speech). In addition, to determine the importance of each region, the number of functional connectivity changes associated with each node was sorted and the result showed dSMC (four significant connections) > vM2 (3) > Broca (2) = dM2 (2) > vSMC (1). Taking together, dSMC-related functional connections were enhanced for covert speech compared to overt speech condition.

4. Discussion

4.1. fNIRS could reveal the hemodynamic response and functional connectivity of cortical activities during the covert and overt speech processing

In this study, fNIRS signals were recorded from the subregions of the motor cortex (dSMC, vSMC, dM2 and vM2) and Broca's area for 13 subjects during the overt and covert speech conditions. By examining the hemodynamic response and the functional connectivity of HbO responses, we found that (a) The bilateral dorsal motor cortex was most significantly activated during the covert speech, whereas the bilateral ventral motor cortex was significantly activated during the overt speech. (b) As a subregion of the motor cortex, SMC showed a significant dominant dorsal response to covert speech condition, whereas a dominant ventral response to overt speech condition. (c) Broca was deactivated during the covert speech but activated during the overt speech. (d) Compared to overt speech, dSMC-related functional connections were significantly enhanced during the covert speech. Taking together, our results suggest that the dorsal motor cortex prefers the covert speech condition, whereas the ventral motor cortex and Broca's area prefer the overt speech condition.

4.2. The HbO concentration change signal as an effective neural marker for revealing cortical activities during the covert and overt speech processing

Previous fNIRS studies showed that the HbO concentration change signals could reflect localized neural processing [3] and could be a promising neural marker [62, 71]. Enhanced neural activity leads to a rise in cerebral blood flow in the arteries in a feed-forward manner. The arteries blood contains higher concentrations of HbO and lower concentrations of HbR. Parallel to this, neural activity leads to an enhanced cerebral metabolic rate of oxygen. Since the amount of HbO replenished by arterial blood is usually higher than consumed by metabolism, the HbO concentration rises in response to enhanced neural activity. This process is known as the 'neurovascular coupling' effect [60, 72]. Our results further validated that HbO signals could reflect cortical activities during speech processing. In this study, the grand-averaged HbO concentration change signals (i.e. grand-averaged across all subjects and all conditions) exhibited a trend of first increasing, then reaching a plateau, and finally decreasing (figure 2(d)), which was consistent with the canonical hemodynamic responses from block-design [73]. The similar changing tendency of HbO concentration changes were revealed under overt or covert speech condition (figure 3). In addition, the HbO and HbR concentration changes in vSMC during the overt speech (figure 3(b)) were also consistent with previous fNIRS findings for the overt speech condition [10], which further confirms the reliability of our hemodynamic response results. Moreover, the trend of negative correlation between HbO and HbR concentration changes signals (figures 2(d) and 3) implied that the information of HbR signals can also be reflected by HbO signals [59, 62]. Nevertheless, the HbO signals had a higher amplitude compared to the HbR signals, which reflected the higher SNR of the HbO signal compared to the HbR signal [10, 71]. Our findings confirm the validity and reliability of the fNIRS hemodynamic response, especially the HbO response, in capturing cortical activity [74] during the overt and covert speech.

4.3. The hemodynamic response in vSMC reflects speech articulation processing

Our results of event-related Hb concentrate on changes suggest bilateral activation of vSMC during the overt speech (figure 4). This is consistent

with iEEG [75], fMRI [76] results, which have reported bilateral vSMC neural response during the overt speech. In addition, several fNIRS studies have also identified vSMC activation during the overt speech [8]. vSMC has been traditionally thought to reflect the face and throat part of the homunculus. Recent studies have further found that vSMC neural activity encodes articulatory kinematic trajectories [77] and can be used to synthesize realistic audible speeches [78]. This suggests that the response of vSMC during the overt speech reflects the function of the motor executive to control the articulators.

In addition, our results also found vSMC activation during the covert speech (figures 5 and 6). This is consistent with iEEG [19], fMRI [25] and other results showing vSMC neural response during the covert speech. Some hypothesis suggests that the brain generates a simulated copy of motor commands during the covert speech without implementing them [25], implying that fewer neural resources are needed compared to overt speech [79]. The covert speech also involves an active inhibitory neural mechanism, which prevents the actual movement of the articulators. This could be the reason why vSMC neural activity during the covert speech is weaker than overt speech (figures 6 and 7).

4.4. The hemodynamic response in dSMC reveals the neural representation of sensorimotor information processing in working memory

Although it is commonly assumed that vSMC is involved in imagined speech processing [25], our neuroimaging evidence suggests that dSMC is also involved in imagined speech processing (figure 5). Our results are in line with several previous neuroimaging evidence, which demonstrated that dSMC is involved in speech production and imagined vocalization. First, Martin *et al* reconstructed the spectrum and modulation features of overt and covert speech by iEEG [80]. They found some electrodes in the left dSMC highly contributed to the reconstruction of covert speech modulation, implying that dSMC activity carries imagined speech information. Second, Stavisky *et al* found that the 'hand-knob area', previously thought to be responsible for controlling arm movements, also showed significant neural firing during the overt speech. In addition, some orofacial movements involving the lips, tongue, and mouth can also cause neural firing in this area [30]. In addition, the accuracy of controlling the BCI cursor while speaking was similar to that of controlling the BCI cursor alone, and the neural firing associated with the speech was attenuated while controlling the BCI cursor [31]. This implies that dSMC is involved in multitask control of hand and orofacial movements, and that information processing for different tasks is independent to some extent. Third, Tian *et al* found that imagined singing showed activation of the left dSMC; and this area was more activated in imagined singing than listening to songs [26]. Perhaps this is because the dSMC encodes more low-level sensorimotor properties of information during imagined singing than higher-order, conceptual properties shared by listening and imagined singing. Our study adds new fNIRS neuroimaging evidence for dSMC involvement in imagined speech processing. Combined with previous neuroimaging evidence, we hypothesize that the neural activity of dSMC is related to low-level sensorimotor information processing of imagined speech.

In addition, dSMC may act as a 'container' for verbal working memory, storing sensorimotor representations during imagined speech. According to Baddeley's working memory model [29], working memory consists of a phonological storage and an active rehearsal system. The active rehearsal system, which executes the phonological loop, refreshes the decaying information in the phonological storage. This system contains a conceptual system, a speech input system (inner ear and inner muscle receptors) and a speech output system (inner mouth) [81]. The concept-output-input-concept cycle is completed by repeating the content over and over again. This process is identical to our experimental task, which required subjects (via overt and covert speech) to repeat the given word continuously. Therefore, we infer that our task involves the working memory system.

What might be the difference in the execution of the active rehearsal between overt speech and covert speech? Phonological information from the conceptual system is sent to the speech output system and converted into muscle motor commands. For overt speech, the phonological output system sends sound to the phonological input system, which generates an actual perceptual and somatosensory input. For covert speech, a simulated movement trajectory is sent, which generates an estimated perceptual and somatosensory input [25, 37]. Because the input of covert speech is not afferent actually from peripheral nerves, it can only rely on sensory–motor representations in working memory to complete the inner rehearsal. Therefore, covert speech refreshes high-level linguistic information in working memory and low-level sensorimotor information during the phonological loop. In addition, we required subjects to imagine the perceived sound and vocal tract motion, forcing them to maintain a clear somatosensory representation. Thus, we suggest that covert speech requires additional sensory–motor representations in working memory than overt speech. This may account for the stronger dSMC activity in covert speech compared to overt speech (figures 6 and 7).

Other evidence demonstrates that dSMC is engaged in the working memory processing of sensory–motor representations. SMC can be involved in mental imagery, working memory, and other cognitive functions [82]. SMC neurons can 'hold

premotor information for short periods' [83], which might be processed by distributed networks instead of discrete representation [83]. dSMC may be part of the SMC distributed working memory network. On the one hand, Müsch *et al* found that dSMC showed sentence-specific activity in listening and silent rehearsal [24]. This demonstrates that the neural activity of dSMC can encode sentences' articulatory and sensory information sequences. On the other hand, dSMC plays a key role in both decoding imagined hand motor gestures [83] and reconstructing the acoustic features of imagined speech [80]. This means that motor gesture information can be decoded from the cortical activity of the dSMC. In short, we provide fNIRS evidence confirming the cortical activity of dSMC in covert speech. We hypothesize that dSMC is involved in the processing of sensorimotor information in working memory.

4.5. The hemodynamic response of M2 might reflect high-level speech motor processing, such as motor planning, expectation and monitoring

A higher trend of mean HbO concentration changes and a dominant trend of dorsal response percentage (figure 6) were also found in dM2 during the covert speech compared to overt speech. dM2 is associated with the expectation of sensorimotor information of covert speech [24] and is involved in verbal and nonverbal working memory tasks [29]. In addition, dM2 also showed some activation during the overt speech (figure 4), which might be related to motor monitoring during the overt speech [28].

A higher trend of mean HbO concentration changes and a dominant trend of ventral response percentage (figure 6) were found in vM2 during the overt speech compared to covert speech. vM2 generated correction signals when uttered speech was inconsistent with expectations [22], implying its involvement in motor monitoring. In addition, vM2 showed some activation during the covert speech (figure 5), which might be related to the inner planning of movement during the covert speech [25].

Furthermore, the hemodynamic response of one channel in left dM2 was stronger during the covert speech than overt speech (figure 7), which may be related to the higher working memory occupancy of imagined speech. However, our region-wise comparison results showed that neural activity on M2 was not significantly different between overt and covert speech (figure 6). This may indicated that M2 engaged in the high-level information processing of speech as a whole, such as phonological code retrieval [84], which is shared between overt and covert speech.

4.6. The function of Broca's area varied between overt and covert speech

Broca was deactivated during the covert speech (figure 5). Decreased hemodynamic response in Broca may be due to speech motor inhibition [32]. Previous fNIRS studies found activation of the Broca during the overt speech in normal subjects and deactivation in people who stutter [85], which may be due to a speech motor inhibition in stutterers, while triggers stronger deactivation [86, 87]. Moreover, Broca activated during the overt speech (figure 4), and the direct comparison of the two speech conditions further confirms that Broca preferred overt speech condition (figures 6 and 7). Previous studies by iEEG [75], fMRI [76], and fNIRS [8, 10] also found activation of bilateral Broca during the overt speech, suggesting that Broca is involved in motor planning by producing phonological words [19, 88] during the processing of overt speech condition [1]. In summary, the deactivation during the covert speech and activation during the overt speech in Broca implies that it is involved in covert and overt speech conditions in different ways. The function in the covert speech condition may be more inclined to motor inhibition, while in the overt speech condition may be more inclined to motor planning.

4.7. The cooperated cortical network with dSMC as hub encodes the sensory–motor information during the covert speech processing

In our fNIRS study, covert speech showed enhanced functional connectivity of dSMC with ventral motor cortex and Broca's area (figure 8). An fMRI study also identified the presence of functional connectivity between dSMC and vSMC, vM2 as well as Broca during the covert speech [16], which was part of the dorsal stream of speech processing [14] for transforming phonological information into motor information. The enhanced functional connectivity during the covert speech may reflect a more detailed information encoding in working memory. An fMRI study found the enhanced functional connectivity between dSMC and Broca during an intonation discrimination task [89], implied that dSMC, in concert with other regions, encoded richer information in working memory. During the covert speech, vSMC, vM2, and Broca were all involved in motor simulation, while dSMC might involved in sensory estimation [24]. Sensory regions interact with multiple motor regions for encoding richer sensory–motor information, and occupy higher working memory. Besides, covert speech showed diminished functional connectivity between dM2 and Broca compared to overt speech (figure 8). An fNIRS study found that the strength of functional connectivity between dM2 and Broca was weaker during silent reading than oral reading [9], and our study found the same effect during the covert speech compared to overt speech. This may indicate that dM2 and Broca corporately encode less detailed motor commands in motor simulation processing during the covert speech, and therefore occupy less working memory.

4.8. Potential contributions to the future fNIRS-based speech decoding BCI

We found a strong hemodynamic response in dSMC and dM2 during the covert speech. In addition, there exists a specific functional connection between dSMC and ventral motor/Broca during the covert speech. These may inspire the channel layout and feature extraction of future fNIRS speech BCIs. There was a lack of fNIRS evidence of the engagement of the dorsal motor cortex for covert speech. Probably because of this, previous fNIRS speech BCIs also lacked coverage of the dorsal motor cortex. Besides, functional connectivity features have been applied in EEG speech BCI [90], while not in fNIRS speech BCI. Our results suggested that cortical activation and functional connectivity of the dorsal motor area are reliable features for idle detection of imagined speech. This is essential for asynchronous BCI (or self-paced BCI). Furthermore, our fNIRS evidence, together with evidence from other neuroimaging modalities [24, 80], suggested that BCI researchers could pay more attention to the potential role of dSMC in decoding imagined speech.

4.9. Limitations and further works

There exist several potential limitations. Firstly, due to the limited number of the current fNIRS sources and detectors, the channel layout of fNIRS could not cover the whole brain. Especially, the lateral sulcus and temporal lobe were lack of courage, which were highly relevant to covert speech in previous fNIRS research [23, 91]. In the future, a whole-brain coverage fNIRS device should be used to explore the neural mechanism of speech imagery and to find the optimal spatial layout for speech decoding BCI. Secondly, although the statistical validity was shown after FDR correction and statistical power test, more subjects can be further recruited to strengthen the confidence of the current findings. For example, it remains unclear whether M2 has significantly different hemodynamic responses between overt and covert speech (figure 6(b)), which could be further validated by recruiting more participants. Thirdly, the current study used Pearson's correlation coefficients as functional connectivity, which could not quantify the strength of directed casual connectivity between different brain regions. In the future, the Granger causality analysis could be used to further reveal the causal relationship between brain regions. Fourthly, although we controlled the audible vocalization and reduced articulator movements, the peripheral muscular activation and minuscule movements (i.e. subvocalization) were somehow inevitable [92]. In the future, EMG could be added to the study to determine whether the neural processing contains an elaborated outcome of motor command [92]. Finally, the activation and the connectivity of dSMC during the covert speech could be reflecting a cognitive process. Since we did not perform covert speech decoding, there is still a lack of fNIRS evidence that whether the dSMC activity carry any information about the actual imagined content. In the future, the fNIRS-based speech decoding BCI should be conducted to further explore the neural coding mechanisms, as well as to provide a more natural, user-friendly way for developing the further BCI.

5. Conclusion

We found that dSMC might be the key node of the speech imagery network, while the functional connection of dSMC with the ventral motor cortex and Broca's area is the key edge of the speech imagery network. Our study provides fNIRS-based neuroimaging evidence for the involvement of the dorsal part of the motor cortex, especially dSMC, during the covert speech. This study is expected to inspire the optimization of fNIRS speech BCI in terms of both channel layout and feature extraction.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

The authors would like to thank all the participants and Mr. Zhuobin Yang. This work was supported by National Natural Science Foundation of China (No. 61906132 to X S and No. 81925020 to D M), Key Project & Team Program of Tianjin City (No. XC202020 to X S), and Tianjin University (No. 2020XRY-0015 to X S). The authors have confirmed that any identifiable participants in this study have given their consent for publication.

ORCID iDs

Xiaopeng Si https://orcid.org/0000-0002-8956-7577
Sicheng Li https://orcid.org/0000-0002-0579-727X

References

[1] Cooney C, Folli R and Coyle D 2018 Neurolinguistics research advancing development of a direct-speech brain-computer interface *iScience* **8** 103–25
[2] Rezazadeh Sereshkeh A, Yousefi R, Wong A T and Chau T 2019 Online classification of imagined speech using functional near-infrared spectroscopy signals *J. Neural Eng.* **16** 016005
[3] Ferrari M and Quaresima V 2012 A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application *NeuroImage* **63** 921–35
[4] Herff C and Schultz T 2016 Automatic speech recognition from neural signals: a focused review *Front. Neurosci.* **10** 1–7

[5] Van Gerven M et al 2009 The brain–computer interface cycle J. Neural Eng. 6 041001
[6] Naseer N and Hong K-S 2015 fNIRS-based brain-computer interfaces: a review Front. Hum. Neurosci. 9 1–15
[7] Yoo S and Lee K M 2013 Articulation-based sound perception in verbal repetition: A functional NIRS study Front. Hum. Neurosci. 7 1–10
[8] Moriai-Izawa A, Dan H, Dan I, Sano T, Oguro K, Yokota H, Tsuzuki D and Watanabe E 2012 Multichannel fNIRS assessment of overt and covert confrontation naming Brain Lang. 121 185–93
[9] Wan N, Hancock A S, Moon T K and Gillam R B 2018 A functional near-infrared spectroscopic investigation of speech production during reading Hum. Brain Mapp. 39 1428–37
[10] Sugiura L, Ojima S, Matsuba-Kurita H, Dan I, Tsuzuki D, Katura T and Hagiwara H 2011 Sound to language: different cortical processing for first and second languages in elementary school children as revealed by a large-scale study using fNIRS Cereb. Cortex 21 2374–93
[11] Fló A, Brusini P, Macagno F, Nespor M, Mehler J and Ferry A L 2019 Newborns are sensitive to multiple cues for word segmentation in continuous speech Dev. Sci. 22 e12802
[12] Hwang H-J, Choi H, Kim J-Y, Chang W-D, Kim D-W, Kim K, Jo S and Im C-H 2016 Toward more intuitive brain–computer interfacing: classification of binary covert intentions using functional near-infrared spectroscopy J. Biomed. Opt. 21 091303
[13] Rezazadeh Sereshkeh A, Yousefi R, Wong A T, Rudzicz F and Chau T 2019 Development of a ternary hybrid fNIRS-EEG brain–computer interface based on imagined speech Brain-Comput. Interfaces 6 128–40
[14] Hickok G and Poeppel D 2007 The cortical organization of speech processing Nat. Rev. Neurosci. 8 393–402
[15] Si X, Zhou W and Hong B 2017 Cooperative cortical network for categorical processing of Chinese lexical tone Proc. Natl Acad. Sci. 114 12303–8
[16] Buchsbaum B R, Olsen R K, Koch P and Berman K F 2005 Human dorsal and ventral auditory streams subserve rehearsal-based and echoic processes during verbal working memory Neuron 48 687–97
[17] Basho S, Palmer E D, Rubio M A, Wulfeck B and Muller R-A 2007 Effects of generation mode in fMRI adaptations of semantic fluency: paced production and overt speech Neuropsychologia 45 1697–706
[18] Doucet G et al 2012 Patterns of hemodynamic low-frequency oscillations in the brain are modulated by the nature of free thought during rest NeuroImage 59 3194–200
[19] Pei X, Leuthardt E C, Gaona C M, Brunner P, Wolpaw J R and Schalk G 2011 Spatiotemporal dynamics of electrocorticographic high gamma activity during overt and covert word repetition NeuroImage 54 2960–72
[20] Perrone-Bertolotti M, Kujala J, Vidal J R, Hamame C M, Ossandon T, Bertrand O, Minotti L, Kahane P, Jerbi K and Lachaux J P 2012 How silent is silent reading? intracerebral evidence for top-down activation of temporal voice areas during reading J. Neurosci. 32 17554–62
[21] Simonyan K, Ackermann H, Chang E F and Greenlee J D 2016 New developments in understanding the complexity of human speech production J. Neurosci. 36 11440–8
[22] Chang E F, Niziolek C A, Knight R T, Nagarajan S S and Houde J F 2013 Human cortical sensorimotor network underlying feedback control of vocal pitch Proc. Natl Acad. Sci. 110 2653–8
[23] Eggebrecht A T, Ferradal S L, Robichaux-Viehoever A, Hassanpour M S, Dehghani H, Snyder A Z, Hershey T and Culver J P 2014 Mapping distributed brain function and networks with diffuse optical tomography Nat. Photon. 8 448–54
[24] Musch K, Himberger K, Tan K M, Valiante T A and Honey C J 2020 Transformation of speech sequences in human sensorimotor circuits Proc. Natl Acad. Sci. USA 117 3203–13
[25] Tian X, Zarate J M and Poeppel D 2016 Mental imagery of speech implicates two mechanisms of perceptual reactivation Cortex 77 1–12
[26] Li Y, Luo H and Tian X 2020 Mental operations in rhythm: motor-to-sensory transformation mediates imagined singing PLoS Biol. 18 1–19
[27] Hanakawa T, Dimyan M A and Hallett M 2008 Motor planning, imagery, and execution in the distributed motor network: a time-course study with functional MRI Cereb. Cortex 18 2775–88
[28] Lima C F, Krishnan S and Scott S K 2016 Roles of supplementary motor areas in auditory processing and auditory imagery Trends Neurosci. 39 527–42
[29] Marvel C L, Morgan O P and Kronemer S I 2019 How the motor system integrates with working memory Neurosci. Biobehav. Rev. 102 184–94
[30] Stavisky S D et al 2019 Neural ensemble dynamics in dorsal motor cortex during speech in people with paralysis eLife 8 1–31
[31] Stavisky S D, Willett F R, Avansino D T, Hochberg L R, Shenoy K V and Henderson J M 2020 Speech-related dorsal motor cortex activity does not interfere with iBCI cursor control J. Neural Eng. 17 016049
[32] Fedorenko E and Blank I A 2020 Broca's area is not a natural kind Trends Cogn. Sci. 24 270–84
[33] Ge J, Peng G, Lyu B, Wang Y, Zhuo Y, Niu Z, Tan L H, Leff A P and Gao J-H 2015 Cross-language differences in the brain network subserving intelligible speech Proc. Natl Acad. Sci. 112 2972–7
[34] Brainard D H 1997 The psychophysics toolbox Spatial Vis. 10 433–6
[35] Schultz T, Wand M, Hueber T, Krusienski D J, Herff C and Brumberg J S 2017 Biosignal-based spoken communication: a survey IEEE/ACM Trans. Audio Speech Lang. Process. 25 2257–71
[36] Scholkmann F, Gerber U, Wolf M and Wolf U 2013 End-tidal CO2: an important parameter for a correct interpretation in functional brain studies using speech tasks NeuroImage 66 71–79
[37] Tian X and Poeppel D 2012 Mental imagery of speech: linking motor and perceptual systems through internal simulation and estimation Front. Hum. Neurosci. 6 1–11
[38] Alderson-Day B and Fernyhough C 2015 Inner speech: development, cognitive functions, phenomenology, and neurobiology Psychol. Bull. 141 931
[39] Tsuzuki D, Jurcak V, Singh A K, Okamoto M, Watanabe E and Dan I 2007 Virtual spatial registration of stand-alone fNIRS data to MNI space NeuroImage 34 1506–18
[40] Holmes C J, Hoge R, Collins L, Woods R, Toga A W and Evans A C 1998 Enhancement of MR images using registration for signal averaging J. Comput. Assist. Tomogr. 22 324–33
[41] Petkov M P, Borsook D, Dubb J, Dan I, Aasted C M, Cooper R J, Ycel M A, Becerra L, Boas D A and Tsuzuki D 2015 Anatomical guidance for functional near-infrared spectroscopy: AtlasViewer tutorial Neurophotonics 2 020801
[42] Okamoto M, Tsuzuki D, Clowney L, Dan H, Singh A K and Dan I 2009 Structural atlas-based spatial registration for functional near-infrared spectroscopy enabling inter-study data integration Clin. Neurophysiol. 120 1320–8
[43] Xia M, Wang J and He Y 2013 BrainNet viewer: a network visualization tool for human brain connectomics PLoS One 8 e68910
[44] Ardila A, Bernal B and Rosselli M 2016 How localized are language brain areas? A review of Brodmann areas involvement in oral language Arch. Clin. Neuropsychol. 31 112–22
[45] Barthas F and Kwan A C 2017 Secondary motor cortex: where 'sensory'meets 'motor'in the rodent frontal cortex Trends Neurosci. 40 181–93

[46] Côté S L, Hamadjida A, Quessy S and Dancause N 2017 Contrasting modulatory effects from the dorsal and ventral premotor cortex on primary motor cortex outputs *J. Neurosci.* **37** 5960–73
[47] Genon S *et al* 2018 The heterogeneity of the left dorsal premotor cortex evidenced by multimodal connectivity-based parcellation and functional characterization *NeuroImage* **170** 400–11
[48] Dumontheil I 2014 Development of abstract thinking during childhood and adolescence: the role of rostrolateral prefrontal cortex *Dev. Cogn. Neurosci.* **10** 57–76
[49] Tachtsidis I and Scholkmann F 2016 False positives and false negatives in functional near-infrared spectroscopy: issues, challenges, and the way forward *Neurophotonics* **3** 031405
[50] Peters S, Lim S B, Louie D R, Yang C L and Eng J J 2020 Passive, yet not inactive: robotic exoskeleton walking increases cortical activation dependent on task *J. Neuroeng. Rehabil.* **17** 1–12
[51] Lloyd-Fox S, Blasi A and Elwell C E 2010 Illuminating the developing brain: the past, present and future of functional near infrared spectroscopy *Neurosci. Biobehav. Rev.* **34** 269–84
[52] Bulgarelli C, Blasi A, De Klerk C C J M, Richards J E, Hamilton A and Southgate V 2019 Fronto-temporoparietal connectivity and self-awareness in 18-month-olds: a resting state fNIRS study *Dev. Cogn. Neurosci.* **38** 100676
[53] Bulgarelli C, Klerk C C J M, Richards J E, Southgate V, Hamilton A and Blasi A 2020 The developmental trajectory of fronto-temporoparietal connectivity as a proxy of the default mode network: a longitudinal fNIRS investigation *Hum. Brain Mapp.* **41** 2717–40
[54] Haeussinger F B, Dresler T, Heinzel S, Schecklmann M, Fallgatter A J and Ehlis A-C 2014 Reconstructing functional near-infrared spectroscopy (fNIRS) signals impaired by extra-cranial confounds: an easy-to-use filter method *NeuroImage* **95** 69–79
[55] Thranitz J, Knauth M, Heldmann M, Küchler J, Münte T F and Royl G 2020 Elevation of intracranial pressure affects the relationship between hemoglobin concentration and neuronal activation in human somatosensory cortex *Hum. Brain Mapp.* **41** 2702–16
[56] Huppert T J, Diamond S G, Franceschini M A and Boas D A 2009 HomER: a review of time-series analysis methods for near-infrared spectroscopy of the brain *Appl. Opt.* **48** D280–98
[57] Pinti P, Scholkmann F, Hamilton A, Burgess P and Tachtsidis I 2019 Current status and issues regarding pre-processing of fNIRS neuroimaging data: an investigation of diverse signal filtering methods within a general linear model framework *Front. Hum. Neurosci.* **12** 1–21
[58] Urquhart E L, Wang X, Liu H, Fadel P J and Alexandrakis G 2020 Differences in net information flow and dynamic connectivity metrics between physically active and inactive subjects measured by functional near-infrared spectroscopy (fNIRS) during a fatiguing handgrip task *Front. Neurosci.* **14** 1–16
[59] Yamada T, Umeyama S and Matsuda K 2012 Separation of fNIRS signals into functional and systemic components based on differences in hemodynamic modalities *PLoS One* **7** e50271
[60] Scholkmann F, Kleiser S, Metz A J, Zimmermann R, Mata Pavia J, Wolf U and Wolf M 2014 A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology *NeuroImage* **85** 6–27
[61] Lawrence R J, Wiggins I M, Anderson C A, Davies-Thompson J and Hartley D E 2018 Cortical correlates of speech intelligibility measured using functional near-infrared spectroscopy (fNIRS) *Hear. Res.* **370** 53–64
[62] Anderson C A, Wiggins I M, Kitterick P T and Hartley D E H 2017 Adaptive benefit of cross-modal plasticity following cochlear implantation in deaf adults *Proc. Natl Acad. Sci.* **114** 10256–61
[63] Wiggins I M, Anderson C A, Kitterick P T and Hartley D E 2016 Speech-evoked activation in adult temporal cortex measured using functional near-infrared spectroscopy (fNIRS): are the measurements reliable? *Hear. Res.* **339** 142–54
[64] Fahr A Nonparametric analysis *The International Encyclopedia of Communication Research Methods* (Hoboken, NJ: Wiley) pp 1–6
[65] Szucs D and Ioannidis J P A 2017 Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature *PLoS Biol.* **15** 1–18
[66] Whitford T J, Jack B N, Pearson D, Griffiths O, Luque D, Harris A W F, Spencer K M and Le Pelley M E 2017 Neurophysiological evidence of efference copies to inner speech *eLife* **6** 1–23
[67] Yang J, Zhang H, Ni J, De Dreu C K W and Ma Y 2020 Within-group synchronization in the prefrontal cortex associates with intergroup conflict *Nat. Neurosci.* **23** 754–60
[68] Niu H and He Y 2014 Resting-state functional brain connectivity: lessons from functional near-infrared spectroscopy *Neuroscientist* **20** 173–88
[69] Zhang Y and Zhu C 2020 Assessing brain networks by resting-state dynamic functional connectivity: an fNIRS-EEG study *Front. Neurosci.* **13** 1–12
[70] Li R, Rui G, Zhao C, Wang C, Fang F and Zhang Y 2020 Functional network alterations in patients with amnestic mild cognitive impairment characterized using functional near-infrared spectroscopy *IEEE Trans. Neural Syst. Rehabil. Eng.* **28** 123–32
[71] Coyle S M, Ward T E and Markham C M 2007 Brain-computer interface using a simplified functional near-infrared spectroscopy system *J. Neural Eng.* **4** 219–26
[72] Buxton R B 2012 Dynamic models of BOLD contrast *NeuroImage* **62** 953–61
[73] Amaro E and Barker G J 2006 Study design in fMRI: basic principles *Brain Cogn.* **60** 220–32
[74] Obrig H and Villringer A 2003 Beyond the visible—imaging the human brain with light *J. Cereb. Blood Flow Metab.* **23** 1–18
[75] Cogan G B, Thesen T, Carlson C, Doyle W, Devinsky O and Pesaran B 2014 Sensory–motor transformations for speech occur bilaterally *Nature* **507** 94–98
[76] Silbert L J, Honey C J, Simony E, Poeppel D and Hasson U 2014 Coupled neural systems underlie the production and comprehension of naturalistic narrative speech *Proc. Natl Acad. Sci. USA* **111** E4687–96
[77] Chartier J, Anumanchipalli G K, Johnson K and Chang E F 2018 Encoding of articulatory kinematic trajectories in human speech sensorimotor cortex *Neuron* **98** 1042–1054.e4
[78] Anumanchipalli G K, Chartier J and Chang E F 2019 Speech synthesis from neural decoding of spoken sentences *Nature* **568** 493–8
[79] Oppenheim G M and Dell G S 2010 Motor movement matters: the flexible abstractness of inner speech *Mem. Cognit.* **38** 1147–60
[80] Martin S, Brunner P, Holdgraf C, Heinze H-J, Crone N E, Rieger J, Schalk G, Knight R T and Pasley B N 2014 Decoding spectrotemporal features of overt and covert speech from the human cortex *Front. Neuroeng.* **7** 1–15
[81] Hickok G, Houde J and Rong F 2011 Sensorimotor integration in speech processing: computational basis and neural organization *Neuron* **69** 407–22
[82] Tomasino B and Gremese M 2016 The cognitive side of M1 *Front. Hum. Neurosci.* **10**
[83] Pilgramm S, De Haas B, Helm F, Zentgraf K, Stark R, Munzert J and Krüger B 2016 Motor imagery of hand actions: decoding the content of motor imagery from brain activity in frontal and parietal motor areas *Hum. Brain Mapp.* **37** 81–93
[84] Indefrey P and Levelt W J 2004 The spatial and temporal signatures of word production components *Cognition* **92** 101–44

[85] Walsh B, Tian F, Tourville J A, Yücel M A, Kuczek T and Bostian A J 2017 Hemodynamics of speech production: An fNIRS investigation of children who stutter *Sci. Rep.* **7**

[86] John J P, Halahalli H N, Vasudev M K, Jayakumar P N and Jain S 2011 Regional brain activation/deactivation during word generation in schizophrenia: fMRI study *Br. J. Psychiatry* **198** 213–22

[87] Chambers C D, Bellgrove M A, Stokes M G, Henderson T R, Garavan H, Robertson H, Morris A P and Mattingley J B 2006 Executive "Brake Failure" following Deactivation of Human Frontal Lobe *J. Cogn. Neurosci.* **18** 444–55

[88] Price C J 2012 A review and synthesis of the first 20years of PET and fMRI studies of heard speech, spoken language and reading *NeuroImage* **62** 816–47

[89] Chien P, Friederici A D, Hartwigsen G and Sammler D 2021 Intonation processing increases task-specific fronto-temporal connectivity in tonal language speakers *Hum. Brain Mapp.* **42** 161–74

[90] Chengaiyan S, Retnapandian A S and Anandan K 2020 Identification of vowels in consonant–vowel–consonant words from speech imagery based EEG signals *Cogn. Neurodyn.* **14** 1–19

[91] Jackson E S, Wijeakumar S, Beal D S, Brown B, Zebrowski P and Spencer J P 2019 A fNIRS investigation of speech planning and execution in adults who stutter *Neuroscience* **406** 73–85

[92] Nalborczyk L, Grandchamp R, Koster E H, Perrone-Bertolotti M and Loevenbruck H 2020 Can we decode phonetic features in inner speech using surface electromyography? *PLoS One* **15** e0233282