Cerebral Hemodynamics in Speech-Related Cortical Areas: Articulation Learning Involves the Inferior Frontal Gyrus, Ventral Sensory-Motor Cortex, and Parietal-Temporal Sylvian Area

Naomi Nakamichi[1], Kouichi Takamoto[2], Hiroshi Nishimaru[2], Kumiko Fujiwara[1], Yusaku Takamura[2], Jumpei Matsumoto[2], Makoto Noguchi[1] and Hisao Nishijo[2]*

[1] Department of Oral and Maxillofacial Surgery, Graduate School of Medicine and Pharmaceutical Sciences, University of Toyama, Toyama, Japan, [2] System Emotional Science, Graduate School of Medicine and Pharmaceutical Sciences, University of Toyama, Toyama, Japan

Although motor training programs have been applied to childhood apraxia of speech (AOS), the neural mechanisms of articulation learning are not well understood. To this aim, we recorded cerebral hemodynamic activity in the left hemisphere of healthy subjects ($n = 15$) during articulation learning. We used near-infrared spectroscopy (NIRS) while articulated voices were recorded and analyzed using spectrograms. The study consisted of two experimental sessions (modified and control sessions) in which participants were asked to repeat the articulation of the syllables "i-chi-ni" with and without an occlusal splint. This splint was used to increase the vertical dimension of occlusion to mimic conditions of articulation disorder. There were more articulation errors in the modified session, but number of errors were decreased in the final half of the modified session; this suggests that articulation learning took place. The hemodynamic NIRS data revealed significant activation during articulation in the frontal, parietal, and temporal cortices. These areas are involved in phonological processing and articulation planning and execution, and included the following areas: (i) the ventral sensory-motor cortex (vSMC), including the Rolandic operculum, precentral gyrus, and postcentral gyrus, (ii) the dorsal sensory-motor cortex, including the precentral and postcentral gyri, (iii) the opercular part of the inferior frontal gyrus (IFGoperc), (iv) the temporal cortex, including the superior temporal gyrus, and (v) the inferior parietal lobe (IPL), including the supramarginal and angular gyri. The posterior Sylvian fissure at the parietal–temporal boundary (area Spt) was selectively activated in the modified session. Furthermore, hemodynamic activity in the IFGoperc and vSMC was increased in the final half of the modified session compared with its initial half, and negatively correlated with articulation errors during

articulation learning in the modified session. The present results suggest an essential role of the frontal regions, including the IFGoperc and vSMC, in articulation learning, with sensory feedback through area Spt and the IPL. The present study provides clues to the underlying pathology and treatment of childhood apraxia of speech.

Keywords: articulation learning, cerebral hemodynamics, inferior frontal gyrus, ventral sensory-motor cortex, area Spt, near-infrared spectroscopy

INTRODUCTION

Producing speech is unique to humans, and is controlled by distributed interactive neural networks of cortical and subcortical structures. These speech-related networks are known to include the posterior inferior frontal gyrus (Broca's area), ventral sensory-motor cortex (vSMC), supplementary motor area, superior temporal gyrus (STG) including the auditory cortex, left posterior planum temporale (area Spt), anterior insula, basal ganglia, thalamus, and cerebellum (1–3). These neural networks, and particularly the vSMC, underlie the precise movement control of different parts of the vocal tract (i.e., the articulators, such as the tongue and lips) as well as the larynx and thoracic respiration, and do so by controlling over 100 muscles (4–6). Despite the high complexity of speech motor control, most children, but not non-human primates, can effortlessly master the skill to speak (7, 8).

Speech sound disorders include deficits in articulation (motor-based speech production) and/or phonology (knowledge and use of the speech sounds and sound patterns of language) (9), and can disturb the intelligibility of speech. The prevalence of speech sound disorders in young children has been reported to be 2–25% (9–11). In particular, apraxia of speech (AOS), which is a sub-categorical sound speech disorder, has been ascribed to brain deficits in speech motor planning/programming (with intact speech-related muscles) (12, 13), by which linguistic/phonological code is converted into spatially and temporally coordinated patterns of muscle contractions for speech production (14). To promote reorganization and plasticity in the articulation-related brain areas, various behavioral intervention programs have been applied to treat speech sound disorders (15). For children with AOS, motor training programs with the same principles as those for limb training have been applied, and have been shown to be effective (16, 17). These findings suggest that training-induced improvements of articulation might be attributed to changes in articulation-related brain areas. However, few previous imaging studies have investigated brain activity changes during actual articulation learning.

The cortical loci for articular learning remain largely unknown. However, it has been suggested that the primary motor cortex is involved in the motor learning of extremities (18–20). Furthermore, conceptual and computational models of speech production suggest that articulation is controlled online using peripheral feedback signals by way of the temporal and parietal cortices (21, 22). Based on these findings, we hypothesized that neural circuits consisting of the vSMC, temporal cortex, and parietal cortex might be involved in articular learning.

To investigate the cortical neural mechanisms of articulation learning, we recorded cerebral hemodynamic activity in healthy subjects who performed articulation learning. Recordings were made using near-infrared spectroscopy (NIRS), and we analyzed the relationships between changes in cerebral hemodynamic activity and articulation performance. NIRS is a non-invasive neuroimaging technique that can measure changes in oxygenated-hemoglobin (Oxy-Hb), deoxygenated-hemoglobin (Deoxy-Hb), and total hemoglobin (Total-Hb) in the cerebral cortex that are associated with local cortical neuronal activity (23). Although NIRS cannot detect subcortical activities, NIRS can be applied with less body and head restriction in a relatively larger space compared with the other imaging methods such as functional magnetic resonance imaging (fMRI) and positron emission tomography (PET), and less sensitive to artifacts induced by mouth movements. Thus, NIRS allows us to measure brain activity under conditions similar to actual clinical environments; we also took simultaneous acoustic measurements of produced sounds during the articulation learning.

METHODS

Subjects

A total of 15 healthy subjects participated in the present study (mean age ± SE: 25.7 ± 1.6 years; 7 male and 8 female). Inclusion criteria were as follows: (1) right-handed according to the Japanese version of the FLANDERS handedness questionnaire (24), (2) native Japanese speaker, (3) no abnormality in dental occlusion, (4) no speech sound disorder, and (5) no experience with articulation training. All subjects were treated in strict compliance with the Declaration of Helsinki and the United States Code of Federal Regulations for the protection of human participants. The experiments were conducted with the full written consent of each subject, and the protocol was approved by the ethical committee of the University of Toyama.

General Experimental Procedure

The subjects sat in a comfortable chair and were fitted with a NIRS head cap. The study consisted of two experimental sessions, which were as follows: (1) a modified condition session that involved articulation with an interocclusal splint, which was used to increase the vertical dimension of occlusion to mimic conditions of articulation disorder, and (2) a control condition session without an interocclusal splint. Each experimental session had 10 cycles; in each cycle, subjects were instructed to repeat the articulation of a Japanese shout "i-chi-ni" consisting of 3 syllables at an approximate rate of one production of the shout per second for 10 s, followed by a resting period of 60 s. The order

of the sessions was randomized and counterbalanced across the subjects. Both sessions were conducted on the same day with a 10 min interval between the two sessions.

Characteristics of the Speech Sound Used in the Study

In "i-chi-ni," the articulatory point of the consonants of both "chi" and "ni" is the alveolar ridge, but their articulation is different; "chi" is an affricate, while "ni" is a nasal sound. The vowel included in the syllables "i-chi-ni" is "I;" in Japanese, "i" is a front vowel that is produced when the frontal surface of the tongue approaches the hard palate to the very limit where no affricate sounds occur, and is generated when the tongue is positioned at the highest position among the basic Japanese vowels. Therefore, changes in the vertical dimension of occlusion are likely to cause vocalization errors of the vowel "i" and words that include "i."

In clinical cases where the vertical dimension of occlusion is increased by wearing a dental prosthesis or by tongue resection, it is difficult to elevate the tongue surface, and articulation of the vowel "i" and syllables containing "i" is prone to impairment. Second, utterance of "i" could minimize activity of the temporal muscles when measuring brain activity with NIRS, because, of all Japanese vowels, this vowel requires the narrowest opening of the width of the mouth, and requires less mandibular movement. Third, "i-chi-ni," which is Japanese for "one two," is used frequently in everyday speech. Therefore, it is easy for healthy people to articulate these syllables in a natural condition, and cognitive load other than articulatory learning is supposed to be low during conditions under which mouth movements are disturbed by an artificial attachment to the teeth.

Interocclusal Splint

To increase the vertical dimension of occlusion by 8 mm, an interocclusal splint was attached to the maxillary teeth in the present study. The body of the device was made of plastic thermoforming sheet (ERKODENT, Germany), and individually made according to the teeth size, shape, and alignment of each subject. The interocclusal splint was made in the dental clinic of the university hospital before the experiments.

Assessment of the Articulated Sounds

Articulated sounds were recorded using a microphone (AT9942, Audio-Technica, Inc., Tokyo) and digitally stored on a computer hard disk at a sampling rate of 44.1 kHz with 16-bit precision. Recorded and digitized sound data were converted into spectrograms and analyzed using a commercial software (Acoustic Core 8; Arcadia, Inc., Osaka). Three independent raters, who were clinically experienced speech-language-hearing therapists, inspected the spectrograms of the articulated sounds. The raters assessed deficient voice bars on the spectrogram of [i], which was defined as vowel reduction (error). The error rates (number of errors/shout) in individual cycles in the control and modified sessions were computed.

Concordance of mean error rates (numbers of errors/shout) across 10 cycles of articulation in the control and modified

sessions across the three raters was assessed using an intraclass correlation coefficient (ICC). ICC estimates and their 95% confidence intervals were calculated based on a mean-rating (number of raters = 3), absolute-agreement, and two-way random-effects model (25). Concordance was evaluated as follows: poor reliability if ICC < 0.5, moderate reliability if ICC ranged between 0.5 and 0.75, good if ICC ranged between 0.75 and 0.9, and excellent reliability if ICC > 0.9 (25). Since ICC was larger than 0.8 in the present study (see Results), the mean error rate (number of errors/shout) was computed across the three raters in each cycle in each subject, and used for later analyses.

To assess the effects of articulation learning, 10 cycles were divided into early and late phases consisting of 5 initial and 5 final cycles, respectively. The mean error rate (number of errors/shout) across 5 cycles in each phase in each experimental session was analyzed using a repeated-measures two-way analysis of variance (ANOVA).
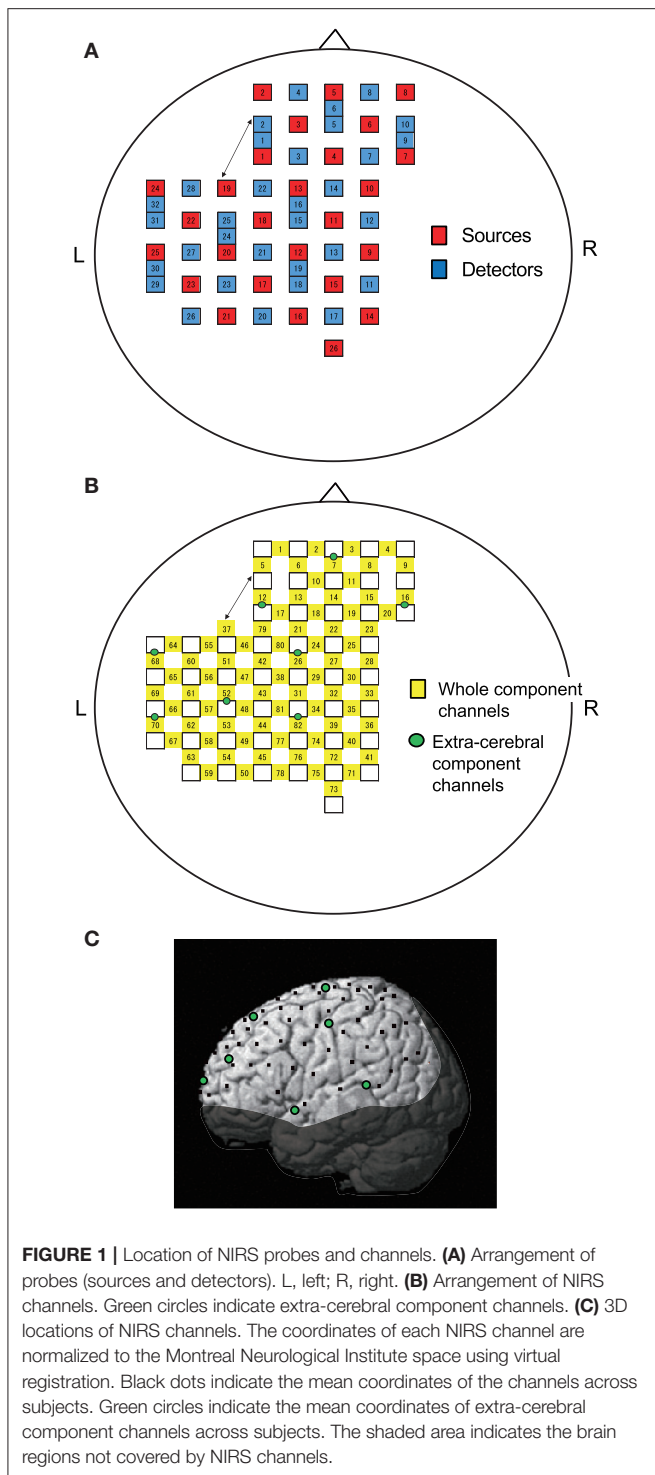
Hemodynamic Measurement and Analysis

Two NIRS systems (OMM3000, Shimadzu Inc., Kyoto) were used to measure cerebral hemodynamics. Twenty-six light-source probes and thirty-two light-detector probes were placed on the head cap (**Figure 1A**). The bottom horizontal line of the probes was placed on the Fp1–Fp2 line according to the 10–20 EEG system. Three different wavelengths (708, 805, and 830 nm) with a pulse width of 5 ms were used to detect hemodynamic responses. The mean total irradiation power was <1 mW. Changes in Hb concentration (Oxy-Hb, Deoxy-Hb, and Total-Hb [Oxy Hb + Deoxy Hb]) were estimated based on a modified Lambert-Beer law (26, 27).

Information included in hemodynamic signals depends on probe distance between light sources and detectors (28–30). The hemodynamic signals from probes with a long distance (more than 3 cm) include not only cerebral (brain) components, but also extra-cerebral (scalp, skull, cerebrospinal fluid) components, while signals measured by probes with a short distance (<1.5 cm) mainly reflect extra-cerebral components. In the present study, multi-distance probe arrangement was applied to remove artifacts and extract cerebral hemodynamics from the whole hemodynamic responses that included both extra-cerebral and cerebral components.

The 22 detector probes were positioned 3 cm from source probes. The midpoints between source and detector probes and signals from those probes were called "whole component channels" and "whole signals," resulting in a total of 82 channels and corresponding signals. Another 8 detector probes were positioned 1.5 cm from source probes, and the midpoints between the probes and signals from those probes were called "extra-cerebral component channels" and "extra-cerebral signals," resulting in a total of 8 channels and corresponding signals (**Figure 1B**). After recording, 3D locations of the NIRS probes were measured using a Digitizer (Real NeuroTechnology Co. Ltd., Japan) with reference to the nasion and bilateral external auditory meatus.

To identify the anatomical locations of the NIRS channels in each subject, the 3D locations of the NIRS probes and channels in

FIGURE 1 | Location of NIRS probes and channels. **(A)** Arrangement of probes (sources and detectors). L, left; R, right. **(B)** Arrangement of NIRS channels. Green circles indicate extra-cerebral component channels. **(C)** 3D locations of NIRS channels. The coordinates of each NIRS channel are normalized to the Montreal Neurological Institute space using virtual registration. Black dots indicate the mean coordinates of the channels across subjects. Green circles indicate the mean coordinates of extra-cerebral component channels across subjects. The shaded area indicates the brain regions not covered by NIRS channels.

each subject were spatially normalized to a standard coordinate system using the software; the coordinates for each NIRS channel were normalized to Montreal Neurological Institute (MNI) space using virtual registration (31) (**Figure 1C**). We then identified the anatomical locations of the NIRS channels of each subject using MRIcro software (www.mricro.com, version 1.4).
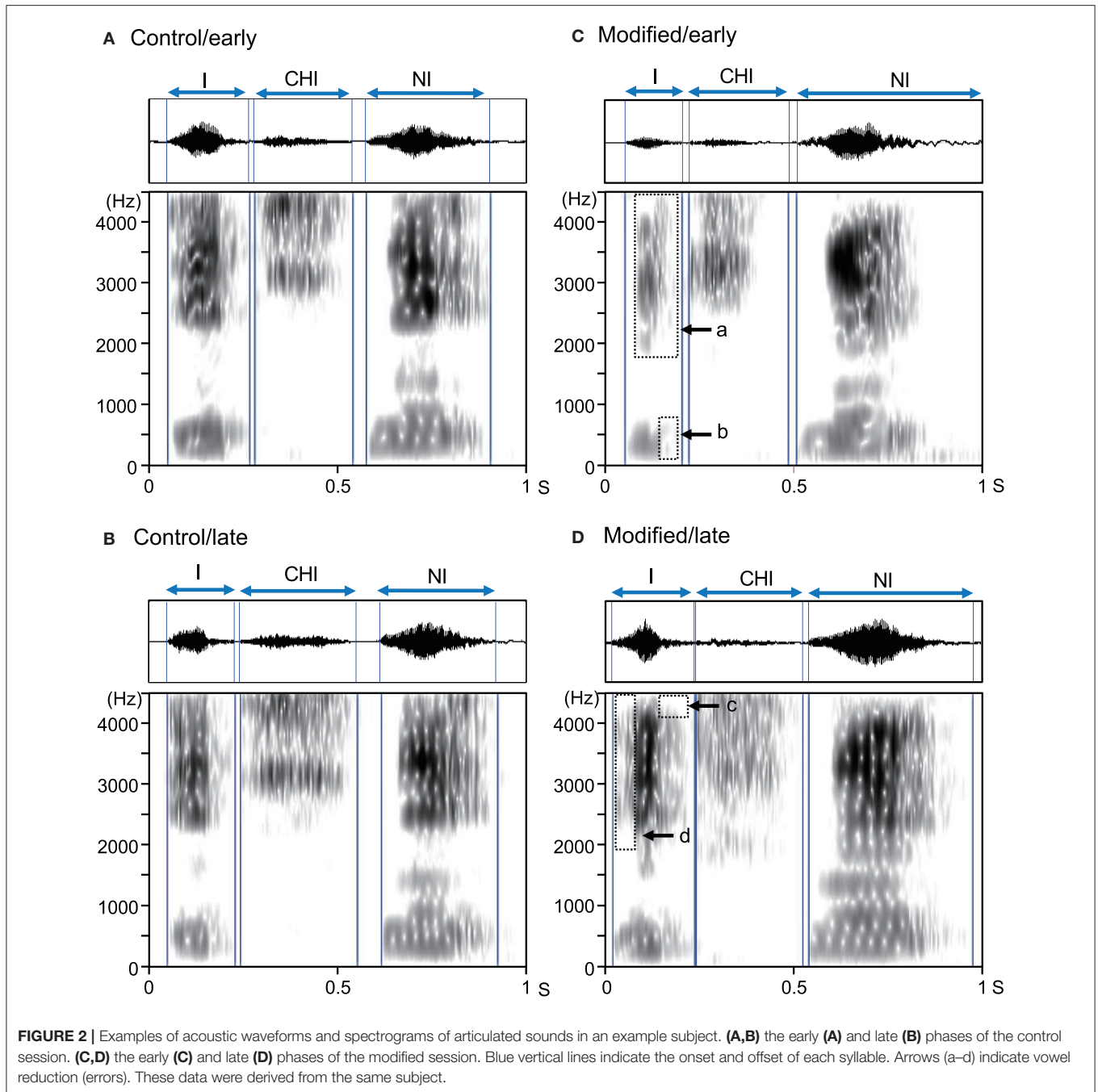
Analysis of NIRS Data

To estimate the cerebral component of NIRS signals (Oxy-Hb, Deoxy-Hb, and Total-Hb), simple-subtraction methods were applied (32); the cerebral hemodynamic activity was calculated by subtraction of the extra-cerebral signals, located nearest to corresponding whole signals, from the whole signals. The subtracted NIRS signals were filtered with a bandpass filter (0.01–0.1 Hz) to remove long-term baseline drift and physiological noise induced by cardiac or respiratory activity (33, 34).

To analyze the time course of hemodynamic responses (Oxy-Hb, Deoxy-Hb, and Total-Hb) during articulation, 10 cycles were divided into early- and late-phase blocks consisting of 5 initial and 5 final cycles, respectively. The NIRS data were summed and averaged for the 5 cycles in each block. The averaged responses were corrected for baseline activity from −5 to 0 s before articulation.

To assess cortical activity, we analyzed Oxy-Hb data, since previous studies have reported that Oxy-Hb is correlated with fMRI BOLD signals (35, 36), and that Oxy-Hb may be the most consistent parameter for cortical activity (37). The cerebral Oxy-Hb data were analyzed based on a mass univariate general linear model (GLM) using NIRS SPM software (statistical parametric mapping: https://www.nitrc.org/projects/nirs_spm/, version 4.1) (38). To analyze the effects of articulation learning on cerebral hemodynamic activity, 10 cycles were divided into early- and late-phase blocks consisting of 5 initial and 5 final cycles, respectively. In the group analysis, the SPM t-statistic maps that were superimposed on the standardized brain according to the MNI coordinate system were generated based on two contrasts, as follows: task (articulation)-related cerebral hemodynamic activity in the early and late blocks, and learning-related cerebral hemodynamic activity between the two blocks. Thus, the contrast maps were as follows: articulation-related images in the early phase, articulation-related images in the late phase, and learning-related contrast images between the late and early phases in the control session, and the same maps in the modified session. Statistical significance was set at an uncorrected threshold of $P < 0.05$.

We also analyzed the correlation between hemodynamic activity during articulation and the error rate (mean number of errors/shout) in individual cycles in the modified session using simple regression analysis. First, the seven brain regions that showed significant activation in the group analysis of comparison between the early and late phases in the modified session were selected as regions of interest (ROIs). The anatomical location and MNI coordinates [(X, Y, Z) mm] of each ROI were as follows: ROI 1 {left opercular part of the inferior frontal gyrus, L-IFGoperc [MNI coordinates: (−60, 15, 23) mm]}, ROI 2 {L-IFGoperc [MNI coordinates: (−60, 8, 10) mm]}, ROI 3 {left precentral gyrus, L-PreCG [MNI coordinates: (−60, 3, 26) mm]}, ROI 4 {left Rolandic operculum, L-ROL [MNI coordinates: (−60, −3, 12) mm]}, ROI 5 {left postcentral gyrus, L-PoCG [MNI coordinates: (−60, −6, 26) mm]}, ROI 6 {L-PoCG [MNI coordinates: (−60, −10, 22) mm]}, and ROI 7 {left supramarginal gyrus, L-SMG [MNI coordinates: (−60, −30, 35) mm]}. Second, T-values of the group analysis in each ROI in individual cycles

FIGURE 2 | Examples of acoustic waveforms and spectrograms of articulated sounds in an example subject. **(A,B)** the early **(A)** and late **(B)** phases of the control session. **(C,D)** the early **(C)** and late **(D)** phases of the modified session. Blue vertical lines indicate the onset and offset of each syllable. Arrows (a–d) indicate vowel reduction (errors). These data were derived from the same subject.

were calculated. Finally, in each ROI, linear regression analysis was performed between t-values and the error rates (numbers of errors/shout) across the 10 cycles.
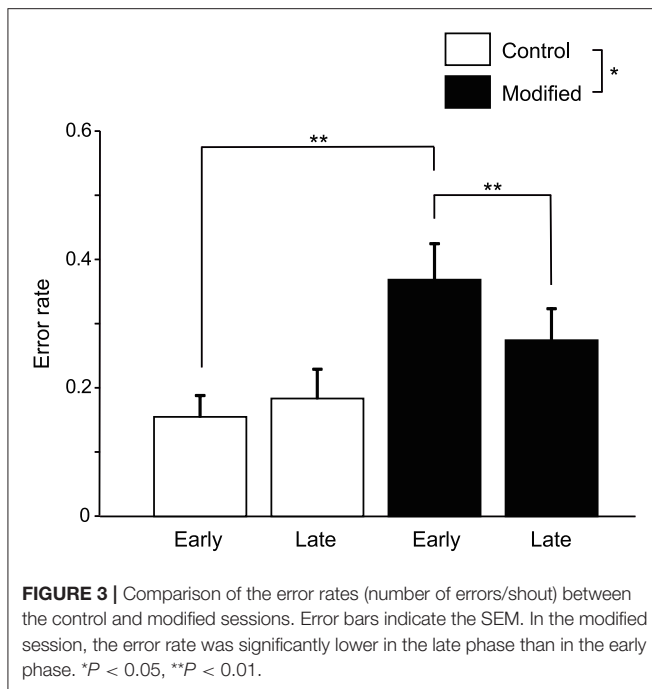
Statistical Analysis

Normality of the data was evaluated using the Shapiro-Wilk test. The homogeneity of variance was evaluated using Levene's test. All statistical analyses were performed using SPSS statistical package version 19.0 (IBM Inc., New York, USA). Values of $P < 0.05$ were considered statistically significant.

RESULTS

Behavioral Changes

Assessment of the all spectrograms of the 15 subjects were consistent across the 3 raters; the ICC value for mean error rate (number of errors/shout) in the control session was 0.94 (95% confidence interval [CI] = 0.86, 0.98), while the ICC value for the mean error rate in the modified session was 0.88 (95% CI = 0.72, 0.96). **Figure 2** shows examples of the acoustic waveforms and spectrograms of articulated sounds in the early and late phases of the control (without an interocclusal splint) and modified (with

FIGURE 3 | Comparison of the error rates (number of errors/shout) between the control and modified sessions. Error bars indicate the SEM. In the modified session, the error rate was significantly lower in the late phase than in the early phase. *$P < 0.05$, **$P < 0.01$.

an interocclusal splint) sessions in one subject. In the control session, no error was observed in the early (**Figure 2A**) and late (**Figure 2B**) phases. However, in the modified session, errors were observed in the early (**Figure 2C**) and late (**Figure 2D**) phases.

Figure 3 shows the comparisons of the error rates (numbers of errors /shout) in the early and late phases of the control and modified sessions. A statistical analysis of the data using a repeated-measures two-way ANOVA with "session" (control vs. modified) and "phase" (early vs. late phases) as factors indicated that there were significant main effects of session [$F_{(1, 14)} = 5.58$, $P < 0.05$] and time [$F_{(1, 14)} = 6.57$, $P < 0.05$], and a significant interaction between session and phase [$F_{(1, 14)} = 13.74$, $P < 0.01$]. *Post-hoc* multiple comparisons indicated that the error rates in the early phase were larger in the modified session than in the control session (Bonferroni test, $P < 0.01$), and that the error rates were lower in the late phase than early phase in the modified session (Bonferroni test, $P < 0.01$).

Hemodynamic Responses

Figure 4 shows examples of the cerebral hemodynamic activity during articulation in the control and modified sessions. Locations of the three cortical channels were as follows: Ch55, L-IFGoperc; Ch60, L-PoCG near the central sulcus; and Ch65, L-STG near the Sylvian fissure (**Figure 4A**). Oxy-Hb and Total-Hb concentrations in the L-IFGoperc, L-PoCG, and L-STG were increased during articulation (**Figure 4B**). There were trends of hemodynamic responses during articulation across the different experimental conditions, as follows. Hemodynamic activity during articulation was larger in the modified session than the control session. Furthermore, hemodynamic activity during articulation was larger in the late than early phases in the

modified session, while it was larger in the early phase than in the late phase in the control session.

Figure 5 shows contrast images resulting from the group analysis based on the GLM with NIRS-SPM. In the early phase of the control session (**Figure 5A**), articulation-related activity was observed in the vSMC, including the L-ROL and L-PoCG, as well as in the dorsal SMC, including the L-Pre/PoCG. Furthermore, activation was observed in the temporal cortex, including the L-STG, and in the inferior parietal lobe (IPL), including the L-SMG and left angular gyrus (L-ANG). In the late phase of the control session (**Figure 5B**), articulation-related activity was less evident; there was activation in the L-IFGoperc, PoCG, L-SMG, and left IPL (L-IPL). The green dotted circles in **Figure 5** indicate the approximate location of the posterior part of the planum temporale (posterior Sylvian fissure at the parietal–temporal boundary, area Spt) based on the data derived from (39). Almost no activation was observed in area Spt, although activation was observed in the L-STG in the control session. The contrast image map between the early and late phases in the control session revealed that only the L-SMG was significantly activated (**Figure 5C**).
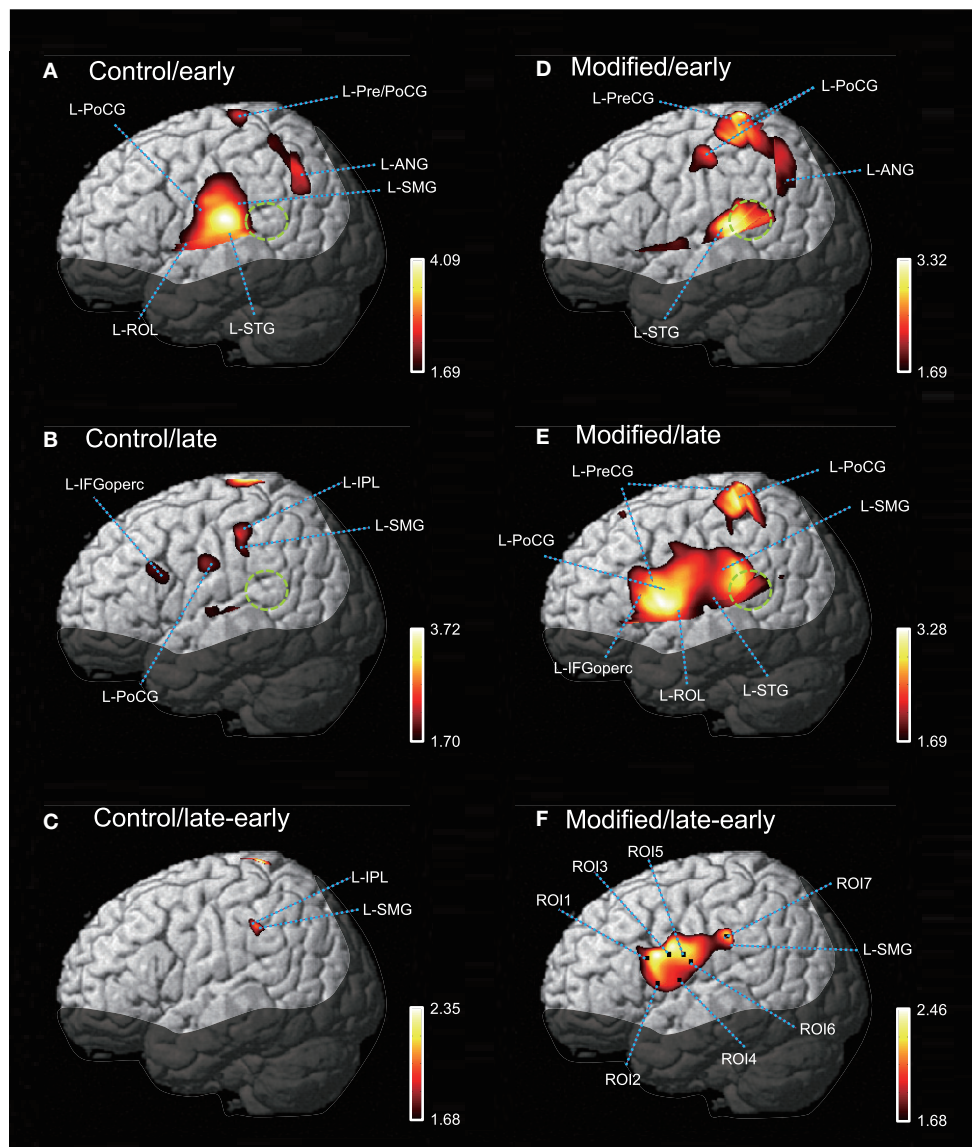
In the early phase of the modified session (**Figure 5D**), activation in the vSMC was less evident; only activation in the dorsal and ventral parts of the SMC was observed. In the dorsal SMC, as well as the L-ANG, a similar articulation-related activation pattern to that in the early phase of the control session was observed in the L-Pre/PoCG and L-ANG. In the temporal cortex, significant activation was observed in the L-STG as well as area Spt (**Figure 5D**). In the late phase of the modified session (**Figure 5E**), wider activation in and around the vSMC was observed in the L-IFGoperc, L-ROL, L-PreCG, L-PoCG, and L-SMG. In the temporal cortex, activation was observed in the L-STG and area Spt (**Figure 5E**). The contrast image map between the early and late phases in the modified session (**Figure 5F**) revealed activation in the L-IFGoperc and vSMC (L-PreCG, L-PoCG, and L-ROL), as well as the L-SMG.

Relationships Between Hemodynamic Responses and Articulation Errors

The contrast image map between the early and late phases in the modified session (**Figure 5F**) indicated significantly greater activation in L-IFGoperc, vSMC, and L-SMG in the late phase than in the early phase (see above). These results suggested that these brain areas might be involved in articulation learning. We investigated this possibility by analyzing the relationship between hemodynamic responses in these areas and the error rates (numbers of errors/shout) across the 10 cycles. **Figure 6** shows the correlation between the articulation-related cortical activation (T-values) in the 6 ROIs in the L-IFGoperc and vSMC and the error rates in the modified session. The 3D locations of the ROIs are shown in **Figure 5F**. There were significant negative correlations between the error rates and hemodynamic responses (T-value) across 10 cycles in ROI 1 (L-IFGoperc) [$F_{(1, 9)} = 15.20$, $P < 0.01$; $r^2 = 0.66$], ROI 2 (L-IFGoperc) [$F_{(1, 9)} = 13.97$, $P < 0.01$; $r^2 = 0.64$], ROI 3 (L-PreCG) [$F_{(1.9)} = 33.21$, $P < 0.01$; $r^2 = 0.81$], ROI 4 (L-ROL) [$F_{(1, 9)} = 33.61$, $P < 0.01$; $r^2 = 0.81$],
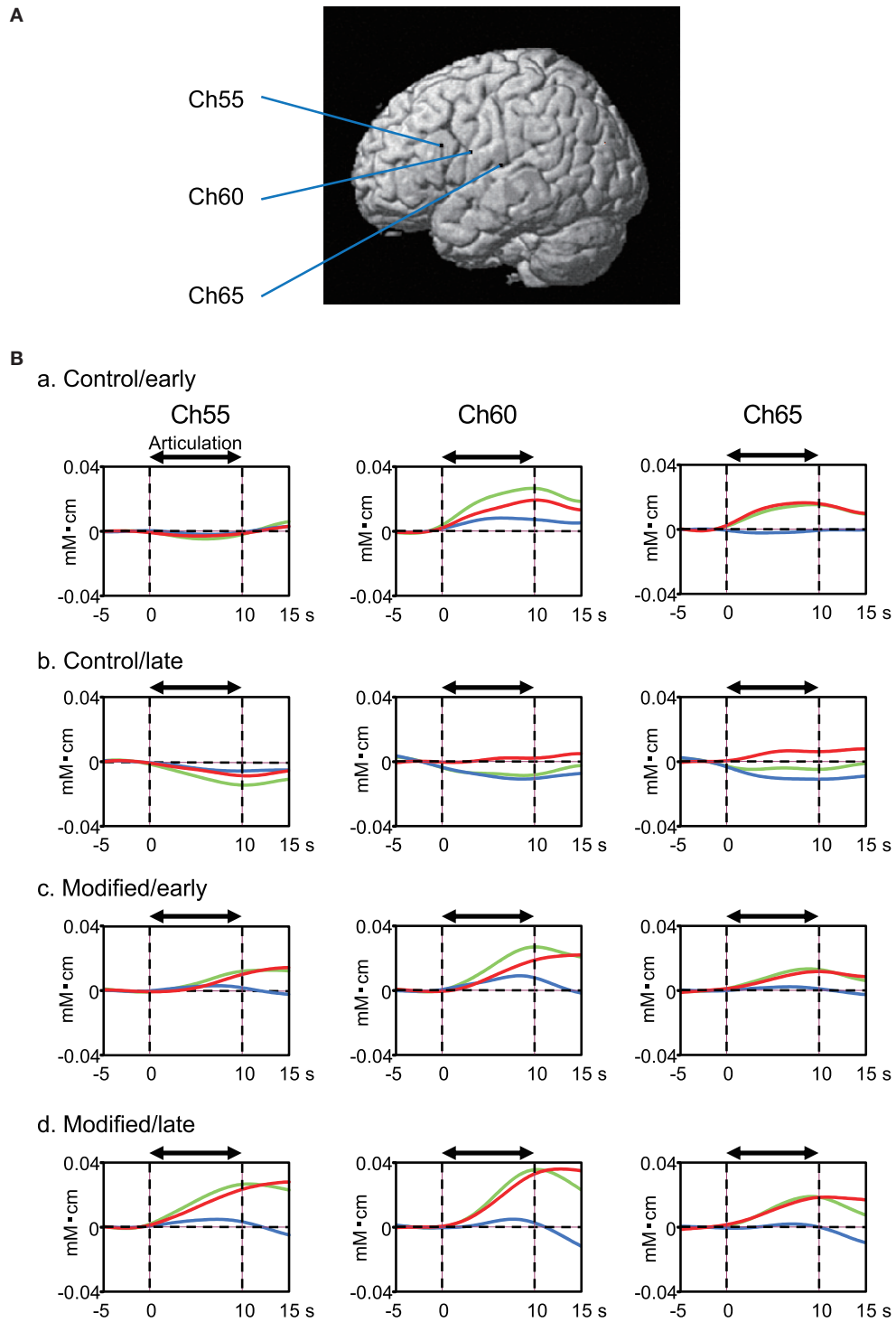
FIGURE 4 | Examples of cerebral hemodynamic activity in the inferior frontal gyrus (Ch55), ventral sensory-motor cortex (Ch60), and superior temporal gyrus (Ch65) during articulation. **(A)** 3D locations of the channels. Ch55, the left opercular part of inferior frontal gyrus [MNI coordinates: (-59, 18, 23) mm]; Ch60, the left post central gyrus [MNI coordinates: (-65, 0, 20) mm]; Ch65, the left superior temporal gyrus [MNI coordinates: (-68, -17, 11) mm]. **(B)** Cerebral hemodynamic activity during articulation in the early (a) and late (b) phases of the control session, and the early (c) and late (d) phases of the modified session. Red, blue, and green lines indicate Oxy-Hb, Deoxy-Hb, and Total-Hb, respectively. The data are derived from the same subject.

FIGURE 5 | The NIRS-SPM T-statistic maps in the control **(A–C)** and modified **(D–F)** sessions. Activation map in the early **(A)** and late **(B)** phases of the control session, and in the early **(D)** and late **(E)** phases of the modified session are shown. Normalized group results of the subtraction of the early phase from the late phase in the control **(C)** and modified **(F)** sessions are also shown. Colored bars indicate the T-values of active voxels. Green dotted circles indicate area Spt (posterior part of the planum temporale). L-ROL, left Rolandic operculum; L-STG, left superior temporal gyrus; L-PoCG, left postcentral gyrus; L-SMG, left supramarginal gyrus; L-IFGoperc, left opercular part of the inferior frontal gyrus; L-PreCG, left precentral gyrus; L-ANG, left angular gyrus.

ROI 5 (L-PoCG) [$F_{(1, 9)} = 28.11$, $P < 0.01$; $r^2 = 0.78$], and ROI 6 (L-PoCG) [$F_{(1, 9)} = 26.67$, $P < 0.01$; $r^2 = 0.77$] (**Figure 6**). However, in the L-SMG, there were no significant correlations between error rates and hemodynamic responses (T-value), even in the ROI with the highest T-value (ROI 7) within the L-SMG [$F_{(1, 9)} = 2.38$, $P = 0.16$; $r^2 = 0.23$].

DISCUSSION

Cortical Activation During Articulation

In the early phase of the control session and the late phase of the modified session, articulation-related activity was observed

in (i) the vSMC, including the L-ROL, L-PreCG, and PoCG, (ii) the dorsal SMC, including the L-Pre/PoCG, (iii) the L-IFGoperc, (iv) the temporal cortex, including the L-STG, and (v) the L-IPL, including the L-SMG and L-ANG. These activated areas are considered to be involved in phonological and articulatory processing.

In the vSMC, including the precentral gyrus (premotor and motor areas), postcentral gyrus (somatosensory area) and subcentral area (Rolandic operculum; both somatosensory and motor-related areas), speech-articulators (i.e., the tongue, lips, jaw, and larynx) are somatotopically represented, and these representations overlap somewhat (4–6, 40). Consistent with
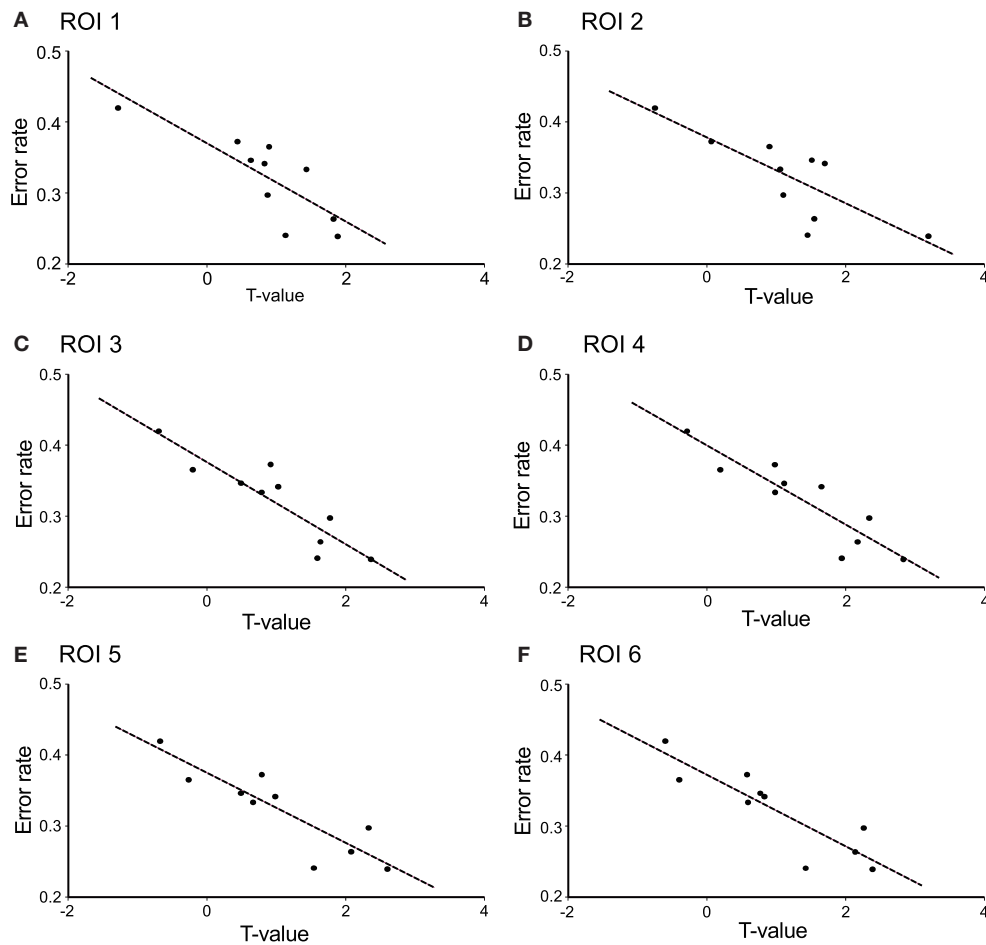
FIGURE 6 | Correlation between cerebral hemodynamic activity (T-values) in the 6 ROIs and syllable production errors (error rates) in the modified session. The hemodynamic activity in all 6 ROIs was significantly negatively correlated with the error rates ($P < 0.01$). **(A)** ROI 1 was the left opercular part of the inferior frontal gyrus; **(B)** ROI 2 was the left opercular part of the inferior frontal gyrus; **(C)** ROI 3 was the left precentral gyrus; **(D)** ROI 4 was the left Rolandic operculum; **(E)** ROI 5 was the left postcentral gyrus; **(F)** ROI 6 was the left postcentral gyrus. The 3D locations of the 6 ROIs are shown in **Figure 5F**.

this previous work, we observed activation in these areas in the early phase of the control session. However, in the late phase of the control session, activation in these areas was less evident, which might be indicative of efficient processing due to adaption in the early phase. The dorsal SMC activation observed in the present study might be involved in respiratory control during syllable production, as has been noted in previous works (1, 41). In the present study, activation of the left posterior part of classic Broca's area (pars opercularis of the left inferior frontal gyrus, L-IFGoperc) was observed during articulation. It has been reported that Broca's area consists of heterogeneously functional areas; its posterior part (L-IFGoperc) is involved in phonological processing during overt speech, while its anterior part (pars triangularis and orbital part of the inferior frontal gyrus) is involved in semantic processing (42–44). The syllables used in the present study were a kind of a shout, and had little semantic meaning, which might induce selective activation in the L-IFGoperc involved in phonological processing.

The STG, including the auditory cortex, was activated in the present study. This region might respond to one's own voice, and

has been reported to be important for the processing of produced phonemes or syllables to send auditory feedback to the motor system for online control of articulation (1, 2). Indeed, subjects tend to compensate their vocal production when auditory feedback is artificially altered by changing formant frequencies, introducing a delay, or masking (45–48). The L-SMG, activated in the present study, has been reported to coactivate with the IFGoperc and vSMC in various phonological processing tasks (49, 50). Furthermore, the SMG has been implicated in proprioceptive processing (51), which suggests that the SMG might also encode movements of the articulatory muscles. These findings suggest that the SMG-frontal motor (i.e., IFGoperc and vSMC) system functions as an "articulatory loop" in the dorsal pathway from the temporal auditory and somatosensory cortices to the frontal cortex through the inferior parietal lobe (SMG), in which phonemic as well as oral somatosensory information is mapped onto motor representations for articulation (52, 53). The L-ANG is also linked to the IFGoperc through the dorsal pathway and has been implicated in phonological processing, such as phoneme discrimination (54).

Area Spt showed a characteristic activation pattern; area Spt was activated in the modified session in which articulation learning was required, while most parts of area Spt were silent in the control session in which no articulation learning was required. Area Spt has been reported to be activated both during auditory perception and covert production of speech (39). Furthermore, brain damage regions near area Spt have been associated with conduction aphasia (55, 56), whereby phonemic errors are more frequent than meaning-based errors (57–59). These findings suggest that area Spt is involved in sensory-motor integration, which is essential for articulation learning (2, 21, 39).

Neural Mechanisms of Overt Articulation Learning

In the present study, the subjects decreased speech production errors during articulation learning in the modified session without any feedback from the experimenters. This suggests that they used their own distorted sounds of the syllables and/or oral somatosensory sensation as feedback signals to improve their articulation. Based on conceptual and computational models of speech production (21, 22), articulation might have been improved using feedback signals in the following potential stages: (i) the sounds (and/or somatosensory sensation) produced by articulation are encoded by their own sensory systems and representations created from auditory (as well as somatosensory) feedback are stored in the short-term memory buffer (area Spt or inferior parietal lobe, including the SMG and ANG), (ii) the feedforward representation of articulatory motor commands are created and compared against these stored representations, and (iii) the mismatch between the two representations is used to improve future articulatory commands. Although there are two (dorsal and ventral) pathways between the temporal auditory and frontal articulatory (IFGoperc and vSMC) systems, the dorsal pathway is essential for articulatory learning (21).

In the dorsal pathway, the SMG and ANG receive inputs from the auditory cortex and have reciprocal connections with the IFGoperc and ventral premotor area in the vSMC (60–62), which is involved in articulatory motor planning (63). Since the SMG functions as a storage house of phonological short-term memory (64), feedback from the auditory system might be transferred to the frontal articulatory system through the SMG to improve articulation. Consistent with this idea, inhibition of the SMG by repetitive transcranial magnetic stimulation (rTMS) has been found to disturb adaptive responses to compensate one's own speech when auditory feedback was altered (65). Area Spt is another candidate area that connects the temporal auditory and frontal articulatory systems (2, 39). Since area Spt was activated selectively in the modified session, during which subjects had to learn new articulation, area Spt might have a more specific role in articulation learning using auditory feedback than the SMG and ANG.

In the present study, hemodynamic activity in ROIs 1–6 (IFGoperc and vSMC) was negatively correlated with articulation error rates. These findings indicate these areas to have a strong association with articulation learning. It has been suggested that the motor learning process by repeating a motor task involves the primary motor cortex, where motor learning is partly mediated through synaptic plasticity, including long-term potentiation (LTP)-like and long-term depression (LTD)-like mechanisms (18–20). The present results revealed an increase in hemodynamic activity in the late phase of the modified session; this suggests that synaptic changes such as LTP might underlie articulation learning in the IFGoperc and vSMC. It has been reported that excitability in speech-related premotor and motor areas was increased when distorted speech sounds or those with noises were presented (66–68). This suggests that the distorted sounds (i.e., sounds with errors) might have triggered synaptic plasticity in the modified session. As far as we know, the present results provide the first evidence of dynamic relative changes in the IFGoperc and vSMC during articulation learning.

On the other hand, in patients with stroke, regions associated with AOS with deficits in speech motor planning/programming (see Introduction) include the left premotor and motor areas (69), suggesting that the premotor and motor areas in the left vSMC are essential for skilled articulatory control. These results are consistent with the present results in which hemodynamic activity in the IFGoperc and vSMC was associated with articulation improvement. A computer simulation study has suggested that childhood AOS is ascribed to developmental deficits in the formation of feedforward programs to control articulators (70). These findings suggest that childhood AOS might be a result of functional deficits in the IFGoperc and vSMC. Thus, the present results provide a neuroscientific basis of articulation learning, which might be disturbed in AOS.

There are some limitations to the present study. First, the spatial resolution of NIRS is less accurate than fMRI. Second, we did not analyze activity in the cerebellum, basal ganglia, or right cerebral cortex, which are involved in speech production and motor learning (see Introduction). Third, Japanese language has some different phonological patterns from other languages as well as some similarities [e.g., (71, 72)]. Therefore, the present results obtained from native Japanese speakers may not be directly applicable to phonological learning in other languages. Further studies are required to (1) generalize our present findings across different language speakers, (2) to clarify the role of brain regions not analyzed in the present study in articulation learning, and (3) to clarify the neurophysiological mechanisms of articulation learning in the IFGoperc and vSMC. Nevertheless, the present results provide clues to the underlying pathology and treatment of AOS.

AUTHOR CONTRIBUTIONS

NN, KT, MN, and HisN designed research. NN, KT, and KF performed research. NN, KT, HirN, JM, YT, and HisN analyzed data. NN, KT, MN, and HisN wrote the manuscript.

ACKNOWLEDGMENTS

REFERENCES

1. Guenther FH, Ghosh SS, Tourville JA. Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain Lang.* (2006) 96:280–301. doi: 10.1016/j.bandl.2005.06.001

2. Hickok G, Houde J, Rong F. Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron* (2011) 69:407–22. doi: 10.1016/j.neuron.2011.01.019

3. Skipper JI, Devlin JT, Lametti DR. The hearing ear is always found close to the speaking tongue: review of the role of the motor system in speech perception. *Brain Lang.* (2017) 164:77–105. doi: 10.1016/j.bandl.2016.10.004

4. Bouchard KE, Mesgarani N, Johnson K, Chang EF. Functional organization of human sensorimotor cortex for speech articulation. *Nature* (2013) 495:327–32. doi: 10.1038/nature11911

5. Conant D, Bouchard KE, Chang EF. Speech map in the human ventral sensory-motor cortex. *Curr Opin Neurobiol.* (2014) 24:63–7. doi: 10.1016/j.conb.2013.08.015

6. Conant DF, Bouchard KE, Leonard MK, Chang EF. Human sensorimotor cortex control of directly measured vocal tract movements during vowel production. *J Neurosci.* (2018) 38:2955–66. doi: 10.1523/JNEUROSCI.2382-17.2018

7. Fitch WT. The evolution of speech: a comparative review. *Trends Cogn Sci.* (2000) 4:258–67. doi: 10.1016/S1364-6613(00)01494-7

8. Kent RD. The uniqueness of speech among motor systems. *Clin Linguist Phon.* (2004) 18:495–505. doi: 10.1080/02699200410001703600

9. Eadie P, Morgan A, Ukoumunne OC, Ttofari Eecen K, Wake M, Reilly S. Speech sound disorder at 4 years: prevalence, comorbidities, and predictors in a community cohort of children. *Dev Med Child Neurol.* (2015) 57:578–84. doi: 10.1111/dmcn.12635

10. Law J, Boyle J, Harris F, Harkness A, Nye C. Prevalence and natural history of primary speech and language delay: findings from a systematic review of the literature. *Int J Lang Commun Disord.* (2000) 35:165–88. doi: 10.1080/136828200247133

11. Wren Y, Miller LL, Peters TJ, Emond A, Roulstone S. Prevalence and predictors of persistent speech sound disorder at eight years old: findings from a population cohort study. *J Speech Lang Hear Res.* (2016) 59:647–73. doi: 10.1044/2015_JSLHR-S-14-0282

12. McNeil MR, Robin DA, Schmidt RA. Apraxia of speech: definition and differential diagnosis. In: McNeil MR, editor. *Clinical Management of Sensorimotor Speech Disorders.* 2nd ed. New York, NY: Thieme (2009). p. 249–68.

13. van der Merwe A. A theoretical framework for the characterization of pathological speech sensorimotor control. In: McNeil MR, editor. *Clinical Management of Sensorimotor Speech Disorders.* 2nd ed. New York, NY: Thieme (2009). p. 1–25.

14. Maas E, Robin DA, Wright DL, Ballard KJ. Motor programming in apraxia of speech. *Brain Lang.* (2008b) 106:107–18. doi: 10.1016/j.bandl.2008.03.004

15. Wren Y, Harding S, Goldbart J, Roulstone S. A systematic review and classification of interventions for speech-sound disorder in preschool children. *Int J Lang Commun Disord.* (2018) 53:446–67. doi: 10.1111/1460-6984.12371

16. Maas E, Robin DA, Austermann Hula SN, Freedman SE, Wulf G, Ballard KJ, et al. Principles of motor learning in treatment of motor speech disorders. *Am J Speech Lang, Pathol.* (2008a) 17:277–98. doi: 10.1044/1058-0360(2008/025)

17. Maas E, Gildersleeve-Neumann C, Jakielski KJ, Stoeckel R. Motor-based intervention protocols in treatment of childhood apraxia of speech (CAS). *Curr Dev Disord Rep.* (2014) 1:197–206. doi: 10.1007/s40474-014-0016-4

18. Rioult-Pedotti MS, Friedman D, Hess G, Donoghue JP. Strengthening of horizontal cortical connections following skill learning. *Nat Neurosci.* (1998) 1:230–4.

19. Muellbacher W, Ziemann U, Wissel J, Dang N, Kofler M, Facchini S, et al. Early consolidation in human primary motor cortex. *Nature* (2002) 415:640–4. doi: 10.1038/nature712

20. Jung P, Ziemann U. Homeostatic and nonhomeostatic modulation of learning in human motor cortex. *J Neurosci.* (2009) 29:5597–604. doi: 10.1523/JNEUROSCI.0222-09.2009

21. Hickok G, Poeppel D. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition* (2004) 92:67–99. doi: 10.1016/j.cognition.2003.10.011

22. Guenther FH, Vladusich T. A neural theory of speech acquisition and production. *J Neurolinguist.* (2012) 25:408–22. doi: 10.1016/j.jneuroling.2009.08.006

23. Ferrari M, Quaresima V. A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *Neuroimage* (2012) 63:921–35. doi: 10.1016/j.neuroimage.2012.03.049

24. Nicholls ME, Thomas NA, Loetscher T, Grimshaw GM. The Flinders Handedness survey (FLANDERS): a brief measure of skilled hand preference. *Cortex* (2013) 49:2914–26. doi: 10.1016/j.cortex.2013.02.002

25. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* (2016) 15:155–63. doi: 10.1016/j.jcm.2016.02.012

26. Seiyama A, Hazeki O, Tamura M. Noninvasive quantitative analysis of blood oxygenation in rat skeletal muscle. *J Biochem.* (1988) 103:419–24. doi: 10.1093/oxfordjournals.jbchem.a122285

27. Wray S, Cope M, Delpy DT, Wyatt JS, Reynolds EO. Characterization of the near infrared absorption spectra of cytochrome aa3 and haemoglobin for the non-invasive monitoring of cerebral oxygenation. *Biochim Biophys Acta* (1988) 933:184–92. doi: 10.1016/0005-2728(88)90069-2

28. Fukui Y, Ajichi Y, Okada E. Monte Carlo prediction of near-infrared light propagation in realistic adult and neonatal head models. *Appl Opt.* (2003) 42:2881–7. doi: 10.1364/AO.42.002881

29. Niederer P, Mudra R, Keller E. Monte Carlo simulation of light propagation in adult brain: influence of tissue blood content and indocyanine green. *Opto Electron Rev.* (2008) 16:124–30. doi: 10.2478/s11772-008-0012-5

30. Ishikuro K, Urakawa S, Takamoto K, Ishikawa A, Ono T, Nishijo H. Cerebral functional imaging using near-infrared spectroscopy during repeated performances of motor rehabilitation tasks tested on healthy subjects. *Front Hum Neurosci.* (2014) 8:292. doi: 10.3389/fnhum.2014.00292

31. Tsuzuki D, Jurcak V, Singh A, Okamoto M, Watanabe E, Dan I. Virtual spatial registration of stand-alone fNIRS data to MNI space. *NeuroImage* (2007) 34:1506–18. doi: 10.1016/j.neuroimage.2006.10.043

32. Schytz HW, Wienecke T, Jensen LT, Selb J, Boas DA, Ashina M. Changes in cerebral blood flow after acetazolamide: an experimental study comparing near-infrared spectroscopy and SPECT. *Eur. J. Neurol.* (2009) 16:461–67. doi: 10.1111/j.1468-1331.2008.02398.x

33. Cordes D, Haughton VM, Arfanakis K, Carew JD, Turski PA, Moritz CH, et al. Frequencies contributing to functional connectivity in the cerebral cortex in "resting-state" data. *AJNR Am J Neuroradiol.* (2001) 22:1326–33.

34. Lu CM, Zhang YJ, Biswal BB, Zang YF, Peng DL, Zhu CZ. Use of fNIRS to assess resting state functional connectivity. *J Neurosci Methods* (2010) 186:242–9. doi: 10.1016/j.jneumeth.2009.11.010

35. Strangman G, Culver JP, Thompson JH, Boas DA. A quantitative comparison of simultaneous BOLD fMRI and NIRS recordings during functional brain activation. *Neuroimage* (2002) 17:719–31. doi: 10.1006/nimg.2002.1227

36. Yamamoto T, Kato T. Paradoxical correlation between signal in functional magnetic resonance imaging and deoxygenated hemoglobin content in capillaries: a new theoretical explanation. *Phys Med Biol.* (2002) 47:1121–41. doi: 10.1088/0031-9155/47/7/309

37. Okamoto M, Matsunami M, Dan H, Kohata T, Kohyama K, Dan I. Prefrontal activity during taste encoding: an fNIRS study. *Neuroimage* (2006) 31:796–806. doi: 10.1016/j.neuroimage.2005.12.021

38. Ye JC, Tak S, Jang KE, Jung J, Jang J. NIRS-SPM: statistical parametric mapping for near-infrared spectroscopy. *Neuroimage* (2009) 44:428–47. doi: 10.1016/j.neuroimage.2008.08.036

39. Hickok G, Okada K, Serences JT. Area Spt in the human planum temporale supports sensory-motor integration for speech processing. *J Neurophysiol.* (2009) 101:2725–32. doi: 10.1152/jn.91099.2008

40. Dichter BK, Breshears JD, Leonard MK, Chang EF. The control of vocal pitch in human laryngeal motor cortex. *Cell* (2018) 174:21–31. doi: 10.1016/j.cell.2018.05.016

41. Takai O, Brown S, Liotti M. Representation of the speech effectors in the human motor cortex: somatotopy or overlap? *Brain Lang.* (2010) 113:39–44. doi: 10.1016/j.bandl.2010.01.008

42. Burton MW, Locasto PC, Krebs-Noble D, Gullapalli RP. A systematic investigation of the functional neuroanatomy of auditory and visual phonological processing. *Neuroimage* (2005) 26:647–61. doi: 10.1016/j.neuroimage.2005.02.024

43. Clos M, Amunts K, Laird AR, Fox PT, Eickhoff SB. Tackling the multifunctional nature of Broca's region meta-analytically: co-activation-based parcellation of area 44. *Neuroimage* (2013) 83:174–88. doi: 10.1016/j.neuroimage.2013.06.041

44. Tate MC, Herbet G, Moritz-Gasser S, Tate JE, Duffau H. Probabilistic map of critical functional regions of the human cerebral cortex: Broca's area revisited. *Brain* (2014) 137(Pt 10):2773–82. doi: 10.1093/brain/awu168

45. Yates AJ. Delayed auditory feedback. *Psychol Bull.* (1963) 60:213–51. doi: 10.1037/h0044155

46. Houde JF, Jordan MI. Sensorimotor adaptation in speech production. *Science* (1998) 279:1213–6. doi: 10.1126/science.279.5354.1213

47. Bauer JJ, Mittal J, Larson CR, Hain TC. Vocal responses to unanticipated perturbations in voice loudness feedback: an automatic mechanism for stabilizing voice amplitude. *J Acoust Soc Am.* (2006) 119:2363–71. doi: 10.1121/1.2173513

48. Purcell DW, Munhall KG. Compensation following real-time manipulation of formants in isolated vowels. *J Acoust Soc Am.* (2006) 119:2288–97. doi: 10.1121/1.2173514

49. Paulesu E, Frith CD, Frackowiak RS. The neural correlates of the verbal component of working memory. *Nature* (1993) 362:342–5.

50. Sliwinska MW, Khadilkar M, Campbell-Ratcliffe J, Quevenco F, Devlin JT. Early and sustained supramarginal gyrus contributions to phonological processing. *Front Psychol.* (2012) 3:161. doi: 10.3389/fpsyg.2012.00161

51. Goble DJ, Coxon JP, Van Impe A, Geurts M, Van Hecke W, Sunaert S, et al. The neural basis of central proprioceptive processing in older versus younger adults: an important sensory role for right putamen. *Hum Brain Mapp.* (2012) 33:895–908. doi: 10.1002/hbm.21257

52. Saur D, Kreher BW, Schnell S, Kümmerer D, Kellmeyer P, Vry MS, et al. Ventral and dorsal pathways for language. *Proc Natl Acad Sci USA.* (2008) 105:18035–40. doi: 10.1073/pnas.0805234105

53. Maldonado IL, Moritz-Gasser S, Duffau H. Does the left superior longitudinal fascicle subserve language semantics? A brain electrostimulation study. *Brain Struct Funct.* (2011) 216:263–74. doi: 10.1007/s00429-011-0309-x

54. Turkeltaub PE, Coslett HB. Localization of sublexical speech perception components. *Brain Lang.* (2010) 114:1–15. doi: 10.1016/j.bandl.2010.03.008

55. Damasio H, Damasio AR. The anatomical basis of conduction aphasia. *Brain* (1980) 103:337–50. doi: 10.1093/brain/103.2.337

56. Damasio H, Damasio AR. Localization of lesions in conduction aphasia. In: Kertesz A, editor. *Localization in Neuropsychology.* San Diego, CA: Academic (1983). p. 231–43.

57. Damasio H. Neuroanatomical correlates of the aphasias. In: Sarno M, editors. *Acquired Aphasia.* San Diego, CA: Academic (1991). p. 45–71.

58. Damasio AR. Aphasia. *N Engl J Med.* (1992) 326:531–9. doi: 10.1056/NEJM199202203260806

59. Hillis AE. Aphasia: progress in the last quarter of a century. *Neurology* (2007) 69:200–13. doi: 10.1212/01.wnl.0000265600.69385.6f

60. Catani M, Jones DK, Ffytche DH. Perisylvian language networks of the human brain. *Ann Neurol.* (2005) 57:8–16. doi: 10.1002/ana.20319

61. Rushworth MF, Behrens TE, Johansen-Berg H. Connection patterns distinguish 3 regions of human parietal cortex. *Cereb Cortex* (2006) 16:1418–30. doi: 10.1093/cercor/bhj079

62. Petrides M, Pandya DN. Distinct parietal and temporal pathways to the homologues of Broca's area in the monkey. *PLoS Biol.* (2009) 7:e1000170. doi: 10.1371/journal.pbio.1000170

63. Price CJ. The anatomy of language: a review of 100 fMRI studies published in 2009. *Ann NY Acad Sci.* (2010) 1191:62–88. doi: 10.1111/j.1749-6632.2010.05444.x

64. Vallar G, Di Betta AM, Silveri MC. The phonological short-term store-rehearsal system: patterns of impairment and neural correlates. *Neuropsychologia* (1997) 35:795–812. doi: 10.1016/S0028-3932(96)00127-3

65. Shum M, Shiller DM, Baum SR, Gracco VL. Sensorimotor integration for speech motor learning involves the inferior parietal cortex. *Eur J Neurosci.* (2011) 34:1817–22. doi: 10.1111/j.1460-9568.2011.07889.x

66. Osnes B, Hugdahl K, Specht K. Effective connectivity analysis demonstrates involvement of premotor cortex during speech perception. *Neuroimage* (2011) 54:2437–45. doi: 10.1016/j.neuroimage.2010.09.078

67. Nuttall HE, Kennedy-Higgins D, Hogan J, Devlin JT, Adank P. The effect of speech distortion on the excitability of articulatory motor cortex. *Neuroimage* (2016) 128:218–26. doi: 10.1016/j.neuroimage.2015.12.038

68. Nuttall HE, Kennedy-Higgins D, Devlin JT, Adank P. The role of hearing ability and speech distortion in the facilitation of articulatory motor cortex. *Neuropsychologia* (2017) 94:13–22. doi: 10.1016/j.neuropsychologia.2016.11.016

69. Graff-Radford J, Jones DT, Strand EA, Rabinstein AA, Duffy JR, Josephs KA. The neuroanatomy of pure apraxia of speech in stroke. *Brain Lang.* (2014) 129:43–6. doi: 10.1016/j.bandl.2014.01.004

70. Terband H, Maassen B. Speech motor development in childhood apraxia of speech: generating testable hypotheses by neurocomputational modeling. *Folia Phoniatr Logop.* (2010) 62:134–42. doi: 10.1159/000287212

71. Beckman ME, Pierrehumbert JB. Intonational structure in Japanese and English. *Phonology Yearbook* (1986) 3:255–309. doi: 10.1017/S095267570000066X

72. Ohata K. Phonological differences between Japanese and english: several potentially problematic areas of pronunciation for Japanese ESL/EFL learners. *Asian EFL J.* (2004) 6:5.