Systems/Circuits

Left Dorsal Speech Stream Components and Their Contribution to Phonological Processing

Takenobu Murakami,[1,2]* Christian A. Kell,[1]* Julia Restle,[1] Yoshikazu Ugawa,[2] and Ulf Ziemann[1,3]

[1]Department of Neurology, Goethe University, Frankfurt am Main, D-60528, Germany, [2]Department of Neurology, Fukushima Medical University, Fukushima, 960-1295, Japan, and [3]Department of Neurology & Stroke, and Hertie Institute for Clinical Brain Research, Eberhard-Karls-University, Tübingen, D-72076, Germany

Models propose an auditory-motor mapping via a left-hemispheric dorsal speech-processing stream, yet its detailed contributions to speech perception and production are unclear. Using fMRI-navigated repetitive transcranial magnetic stimulation (rTMS), we virtually lesioned left dorsal stream components in healthy human subjects and probed the consequences on speech-related facilitation of articulatory motor cortex (M1) excitability, as indexed by increases in motor-evoked potential (MEP) amplitude of a lip muscle, and on speech processing performance in phonological tests. Speech-related MEP facilitation was disrupted by rTMS of the posterior superior temporal sulcus (pSTS), the sylvian parieto-temporal region (SPT), and by double-knock-out but not individual lesioning of pars opercularis of the inferior frontal gyrus (pIFG) and the dorsal premotor cortex (dPMC), and not by rTMS of the ventral speech-processing stream or an occipital control site. RTMS of the dorsal stream but not of the ventral stream or the occipital control site caused deficits specifically in the processing of fast transients of the acoustic speech signal. Performance of syllable and pseudoword repetition correlated with speech-related MEP facilitation, and this relation was abolished with rTMS of pSTS, SPT, and pIFG. Findings provide direct evidence that auditory-motor mapping in the left dorsal stream causes reliable and specific speech-related MEP facilitation in left articulatory M1. The left dorsal stream targets the articulatory M1 through pSTS and SPT constituting essential posterior input regions and parallel via frontal pathways through pIFG and dPMC. Finally, engagement of the left dorsal stream is necessary for processing of fast transients in the auditory signal.

Key words: **articulatory motor cortex; dorsal auditory stream; motor-evoked potential; phonological processing; repetitive transcranial magnetic stimulation; transient virtual lesion**

Introduction

Typically, humans can understand speech reasonably well with their right hemisphere while they produce speech with the left half of their brain. This suggests that the left hemisphere has an advantage over its homolog to extract those features from the acoustic speech signal that are important for learning how to speak (Minagawa-Kawai et al., 2011). A left-lateralized dorsal speech stream has been proposed to map sensory speech representations in the posterior superior temporal lobe via a sensorimotor interface in the sylvian-parieto-temporal region (SPT) in the temporoparietal junction to speech motor representations in the posterior frontal lobe and vice versa (Hickok and Poeppel, 2007; Rauschecker and Scott, 2009). Although the involvement of the dorsal speech stream in speech acquisition is widely accepted, its role in adulthood is still a matter of debate. During speech production, the dorsal stream has been proposed to assist in speech planning by mapping frontal speech motor representations onto temporal representations of their sensory targets to correct intrinsic errors before articulation is performed (Hickok et al., 2011). Such an internal model enables fast enough speech to allow for efficient interpersonal communication without the need of detailed monitoring of external auditory feedback. Internal models have also been proposed to contribute to speech perception in the framework of mirror neurons or the motor theory of speech perception (Liberman and Mattingly, 1985; Fadiga and Craighero, 2003). Although the extreme interpretation of motor involvement in speech perception cannot explain numerous empirical findings (for review, see Hickok, 2009), a modulatory function of the motor cortex in speech perception has been suggested. Speech perception changes the excitability of articulatory motor cortex as measured by transcranial magnetic stimulation (TMS) (Fadiga et al., 2002; Watkins et al., 2003; Murakami et al., 2011), but it is unclear whether this effect is epiphenomenal or causally related to perception. Studies that used repetitive TMS

Received Jan. 19, 2014; revised Oct. 30, 2014; accepted Dec. 3, 2014.

Author contributions: T.M., C.A.K., and U.Z. designed research; T.M. and J.R. performed research; T.M., C.A.K., and J.R. analyzed data; T.M., C.A.K., Y.U., and U.Z. wrote the paper.

T.M. was a fellow of the Alexander von Humboldt Foundation and is currently supported by the Takeda Science and the Kanae Foundation for the Promotion of Medical Science, and the Ministry of Education, Culture, Sports, Science and Technology of Japan. C.A.K. is supported by the German Research Foundation. Financial support did not have any influence on study design, data acquisition, or analysis and publication.

The authors declare no competing financial interests.

*T.M. and C.A.K. contributed equally to this work as first authors.

This article is freely available online through the *J Neurosci* Author Open Choice option.

Correspondence should be addressed to Dr. Ulf Ziemann, Department of Neurology & Stroke, and Hertie Institute for Clinical Brain Research, Eberhard-Karls-University, Hoppe-Seyler-Str. 3, Tübingen, D-72076, Germany. E-mail: ulf.ziemann@uni-tuebingen.de.

DOI:10.1523/JNEUROSCI.0246-14.2015

(rTMS) to induce a transient virtual lesion in speech motor cortex suggest that motor cortex involvement is behaviorally more relevant for speech perception when perceived speech is ambiguous or presented in noise (Meister et al., 2007; Möttönen and Watkins, 2009). In separate studies, three left dorsal speech stream components have been shown to modulate articulatory motor cortex excitability during speech perception: area SPT (Murakami et al., 2012), the dorsal premotor cortex (dPMC) (Meister et al., 2007), and the pars opercularis of the inferior frontal gyrus (pIFG) (Watkins and Paus, 2004; Murakami et al., 2012; Restle et al., 2012). Yet, a systematic study of the specific contributions of the proposed input regions of the dorsal stream to modulation of motor cortex excitability and their contributions to speech processing on the behavioral level is lacking. We thus performed four experiments in which we evaluated the effects of virtual lesions induced by continuous theta burst stimulation (cTBS) (Huang et al., 2005) of the posterior superior temporal sulcus (pSTS), the SPT, dPMC, and pIFG on articulatory motor cortex excitability and on speech-processing performance on the behavioral level.

Materials and Methods

Participants. In total, 24 healthy right-handed native German subjects (mean ± SD age, 25.5 ± 5.1 years, 13 females, Edinburgh Handedness Inventory: 94.8 ± 9.8%) (Oldfield, 1971) participated in this study. Written informed consent was obtained from all subjects before participation. The study was approved by the ethics committee of the Medical Faculty of Goethe-University Frankfurt and conformed to the latest version of the Declaration of Helsinki.

Definition of target regions by fMRI. In all subjects, the individual rTMS target regions were defined on the basis of a prior group fMRI study that focused on audiovisual speech processing with conditions similar to Murakami et al. (2012). In short, fMRI and structural MRI data were acquired on a 3.0-T Siemens Magnetom Allegra System equipped with a standard head coil. fMRI was performed in a block design using a BOLD-sensitive gradient echo planar imaging (EPI) sequence covering the entire brain (repetition time: TR = 2000 ms, echo time: TE = 30 ms, field of view: FOV = 192 mm, flip angle = 90°, 64 × 64 matrix, 30 slices, slice thickness = 2.5 mm, voxel size = 3.0 × 3.0 × 2.5 mm). Eight blocks per condition were randomized and lasted for 30 s each. Each block contained 6 prerecorded German sentences. In the condition of interest, participants passively listened to the presented sentences (80 dB sound pressure level) while viewing 30 s of dynamic visual noise (McConnell and Quinn, 2004). The control condition was listening to 30 s of white noise (80 dB sound pressure level) while viewing 30 s the same visual noise. The 30 s blocks with videos of the lip movements during the production of the prerecorded sentences embedded in white auditory noise were used for defining the target of the visual control region in the middle occipital gyrus (MOG; see below). For coregistration with the functional data, a high-resolution structural T1-weighted image using a magnetic (TR = 2250 ms, TE = 2.6 ms, FOV = 256 mm, flip angle 8°, 256 × 256 matrix, 176 slices, slice thickness = 1 mm, voxel size = 1.0 × 1.0 × 1.0 mm) was acquired. MRI data were standard processed using SPM8 software (Wellcome Department of Imaging Neuroscience, London). EPI volumes were realigned, spatially normalized to the MNI space, and smoothed with an 8 mm at full-width half-maximum Gaussian filter. The preprocessed functional images were analyzed using the standard general linear model approach. Beta maps of the auditory speech condition were contrasted against the control condition in each individual and overlaid on the individual normalized T1 image. Based on prior knowledge on specialized subregions within the dual stream network (Hickok and Poeppel, 2007), we chose the left pIFG, dPMC, SPT, and pSTS as ROIs within the dorsal stream and the left anterior STS/superior temporal gyrus (aSTG) as ROI within the left ventral speech processing stream. The target regions for fMRI-navigated TMS were identified individually as the local individual maxima in the aforementioned anatomical regions at $p < 0.001$, uncorrected (Andoh and Paus, 2011). The individual superior temporal sulcus was chosen as anatomical landmark with the first local maxima anterior of Heschl's gyrus providing coordinates for aSTS/aSTG and with local maxima posterior of Heschl's gyrus for pSTS virtual lesions. Local maxima in the individual temporoparietal junction were designated SPT targets. dPMC targets were located just anterior or within the precentral sulcus at the level of the middle frontal gyrus, whereas pIFG virtual lesions were placed over individual local maxima in the pars opercularis or ventral premotor cortex.

Virtual lesions induced by continuous theta-burst stimulation (cTBS). TBS was delivered by a MagPro X100 magnetic stimulator connected to a 65 mm figure-of-eight coil (MagVenture). The magnetic stimulus had a biphasic waveform with a pulse width of 100 μs. The coil was placed tangentially to the scalp with the handle pointing backwards. Stimulus intensity was set to 80% of active motor threshold as determined over the left motor cortex hot spot for eliciting MEPs in the right first dorsal interosseus muscle, using the relative frequency method (Groppa et al., 2012), and 600 pulses were delivered to the stimulation sites (3 pulses at 50 Hz every 240 ms) (Huang et al., 2005; Murakami et al., 2012). The grand average active motor threshold across all subjects and experiments was 38.9 ± 6.4% (mean ± SD) of maximum stimulator output.

Stimulation sites were identified on each subject's scalp by using an fMRI-guided TMS neuronavigation system (Localite). The individual 3D high-resolution structural T1-weighted image was imported, and a 3D image was standardized to the Talairach coordinate system by defining the AC-PC line and falx cerebri. MNI coordinates were automatically estimated from Talairach ones by using mni2tal (http://imaging.mrc-cbu.cam.ac.uk/downloads/MNI2tal/). Five skull landmarks (nasion, bilateral corners of the eyes, and bilateral preauricular points) and ~150 points on the surface of scalp of each subject were fitted to those of the 3D image. Errors between the subject's scalp and the image were allowed within 1.5 mm. Using the navigation system, the stimulating coil was visually navigated to the stimulation target (individual MNI coordinates; compare Tables 1, 2, and 3) and kept there on the basis of real time feedback of the coil position throughout the cTBS application.

Experiment 1: effects of repetitive TMS on articulatory motor cortex excitability. Experiment 1 tested for contributions of dorsal stream components (left pSTS, SPT, pIFG, dPMC, right pSTS) and the articulatory primary motor cortex (M1) to speech-related facilitation of left articulatory M1 excitability in 8 of the 24 subjects by applying cTBS to these sites (for individual stimulation sites, see Fig. 1*A*; MNI coordinates of individual stimulation sites, Table 1). cTBS of the right pSTS (here individual local maxima in the right STS posterior of Heschl's gyrus were targeted) was performed to study interhemispheric effects, and cTBS of the left articulatory M1 to quantify the direct inhibitory cTBS effect on MEP amplitude. Individual M1 targets were identified as local maxima in the ventral portion of the precentral gyrus or central sulcus. Subjects were seated on a comfortable reclining chair. Before and after cTBS, EMG was recorded from the right orbicularis oris muscle by pairs of Ag-AgCl surface electrodes, bandpass filtered (20–2.000 Hz), then digitized with an analog-to-digital converter (micro1401, CED) at a sampling rate of 5 kHz and stored on a personal computer. As it is difficult to fully voluntarily relax the lip muscles, participants were trained for ~10 min to maintain EMG quiescence before the actual experiments (Murakami et al., 2011). This was achieved by providing participants with continuous high-gain visual (50 μV/Div) and auditory feedback of the EMG activity. TMS for MEP measurements was applied using a Magstim 200 magnetic stimulator and a figure-of-eight coil with external loop diameters of 70 mm (Magstim). Magnetic stimuli had a monophasic waveform. The stimulating coil was held tangentially to the skull with the coil handle pointing backwards and laterally 45° away from the anterior–posterior axis. The center of the coil junction was placed over the motor cortex lip area of the left hemisphere. The "motor hot spot" was determined as the site where TMS consistently elicited the largest MEPs from the right orbicularis oris muscle. This articulatory motor cortex hot spot was located on average 3.1 ± 0.4 cm lateral and 1.9 ± 0.3 cm anterior from the hot spot of the motor cortex hand area (mean ± SD) (Murakami et al., 2011). MEPs were acquired from the right orbicularis oris muscle during listening to prerecorded German declarative sentences (female speaker) or during a control noise condition (listening to white noise) before and

Table 1. MNI coordinates of individual stimulation sites in the first experiment

	Left dPMC			Left pIFG		
	x	*y*	*z*	*x*	*y*	*z*
A	−48	−4	56	−54	4	16
B	−46	−6	50	−48	8	20
C	−48	−6	54	−50	12	16
D	−50	−6	54	−58	0	8
E	−48	−6	54	−52	8	20
F	−50	−4	56	−52	8	20
G	−50	−6	54	−50	2	24
H	−48	4	52	−52	8	20
Average ± SD	−48.5 ± 1.4	−4.3 ± 3.4	53.8 ± 2.0	52.0 ± 3.0	6.3 ± 3.9	18.0 ± 4.7
	Left M1			Left SPT		
	x	*y*	*z*	*x*	*y*	*z*
A	−54	−12	50	−68	−38	26
B	−40	−16	42	−52	−34	18
C	−60	−8	26	−52	−42	26
D	−52	−14	36	−52	−38	22
E	−50	−8	38	−52	−38	26
F	−46	−16	34	−52	−38	32
G	−46	−10	34	−46	−34	24
H	−48	−14	38	−52	−42	26
Average ± SD	−49.5 ± 6.0	−12.3 ± 3.3	37.3 ± 6.9	−53.3 ± 6.3	−38.0 ± 3.0	25.0 ± 4.0
	Left pSTS			Right pSTS		
	x	*y*	*z*	*x*	*y*	*z*
A	−68	−38	18	66	−20	14
B	−54	−36	4	66	−38	14
C	−54	−44	10	64	−38	16
D	−58	−42	8	62	−32	12
E	−50	−42	10	62	−38	10
F	−54	−44	10	60	−36	12
G	−54	−42	8	64	−26	12
H	−56	−46	12	62	−36	8
Average ± SD	−56.0 ± 5.3	−41.8 ± 3.3	10.0 ± 4.0	63.3 ± 2.1	−33.0 ± 6.7	12.3 ± 2.5

Table 2. MNI coordinates of individual stimulation sites in the second experiment

	Left SPT			Left pSTS		
	x	*y*	*z*	*x*	*y*	*z*
A	−56	−38	14	−62	−38	2
B	−46	−40	30	−64	−48	12
C	−66	−36	22	−66	−36	10
D	−54	−44	26	−54	−44	10
E	−54	−42	24	−56	−46	6
F	−54	−42	22	−68	−42	8
G	−62	−50	22	−62	−40	6
H	−52	−42	26	−48	−42	6
I	−60	−50	26	−54	−44	10
J	−58	−34	20	−64	−40	4
K	−52	−40	24	−56	−40	10
L	−48	−38	26	−58	−40	14
M	−48	−46	18	−50	−46	2
N	−56	−44	26	−60	−40	4
O	−50	−44	24	−58	−40	10
P	−50	−40	24	−56	−38	8
Q	−52	−38	22	−58	−42	8
R	−52	−38	32	−54	−44	10
S	−68	−38	26	−68	−38	18
Average ± SD	−54.6 ± 6.0	−41.3 ± 4.3	23.9 ± 4.0	−58.7 ± 5.7	−41.5 ± 3.2	8.3 ± 4.0
	Left pIFG			Left MOG		
	x	*y*	*z*	*x*	*y*	*z*
A	−52	8	20	−44	−74	6
B	−52	8	20	−44	−74	6
C	−50	14	14	−42	−88	10
D	−60	8	20	−44	−74	6
E	−54	8	20	−50	−82	10
F	−62	8	14	−44	−74	6
G	−62	8	20	−50	−80	8
H	−58	16	14	−44	−74	6
I	−52	2	20	−46	−72	4
J	−52	8	22	−44	−68	6
K	−52	8	20	−48	−84	12
L	−46	10	14	−58	−68	12
M	−52	8	20	−44	−74	4
N	−52	8	12	−44	−74	−8
O	−52	8	20	−44	−74	0
P	−52	8	18	−50	−78	6
Q	−58	0	8	−50	−76	8
R	−52	8	20	−50	−76	4
S	−54	4	16	−50	−76	6
Average ± SD	−53.9 ± 4.2	7.9 ± 3.5	17.5 ± 3.8	−46.8 ± 4.0	−75.8 ± 4.9	5.9 ± 4.4

after cTBS. The stimulus intensity was adjusted to elicit MEP peak-to-peak amplitudes in the orbicularis oris muscle of on average 0.3 mV during the control condition at baseline. This intensity was on average (across all subjects and experiments) 62.7 ± 12.3% (mean ± SD) of maximum stimulator output. Additionally, facilitating intermittent TBS (iTBS, 600 pulses at 80% active motor threshold, 3 pulses at 50 Hz every 240 ms for trains of 2 s every 10 s) (Huang et al., 2005; Restle et al., 2012) was applied over the individual left pSTS, and effects were compared with the virtual lesion effects induced by cTBS over the left pSTS. The iTBS condition was added to document the specificity of the interventional TBS protocol and to rule out that nonspecific effects (e.g., on attention) could account for the observed cTBS effects (see below). The order of cTBS sites and the iTBS session was randomized across subjects, and sessions in a given individual were at least 1 week apart to avoid carryover effects.

Given the close adjacency of the stimulation sites over pSTS and SPT (see Fig. 1*A*), it is not certain that selective stimulation of either of these two targets without spillover to nontargeted area has been achieved. Other studies revealed that this is possible in the temporoparietal junction (Kelly et al., 2014) and in the posterior parietal cortex (Koch et al., 2007) by demonstrating dissociable TMS effects on behavior or excitability. However, in our experiments, we expected (and indeed observed; see Results) largely identical rather than dissociable cTBS effects when targeting pSTS and SPT. This, purely on logical grounds, precludes a firm conclusion as to whether or not pSTS and SPT have been stimulated separately.

Absolute MEP amplitudes before and after cTBS were compared using a three-way repeated-measures ANOVA with experimental condition (2 levels: listening to speech vs listening to white noise), stimulus site (6 levels: left pSTS, right pSTS, left SPT, left pIFG, left dPMC, left M1), and time (2 levels: pre- vs post-cTBS) as within-subject factors. For comparison of iTBS with cTBS over left pSTS, another three-way repeated-measures ANOVA with intervention (2 levels: cTBS vs iTBS), experimental condition (2 levels), and time (2 levels) as within-subject factors was conducted. Whenever there was a significant interaction, *post hoc* paired *t* tests were performed using Fisher's protected least significant difference (Fisher's PLSD) test.

Experiment 2: temporal evolution and duration of cTBS effects. Five of the eight subjects who had participated in Experiment 1 joined Experiment 2 to test the temporal evolution and duration of cTBS effects over the left pSTS. MEP amplitudes were recorded from the right orbicularis oris muscle in the listening to speech condition versus control noise condition at five time points: pre-cTBS, and post-cTBS at 0–5, 10–15, 20–25, and 30–35 min. A two-way repeated-measures ANOVA with experimental condition (2 levels) and time (5 levels) was conducted. *Post*

Table 3. MNI coordinates of individual stimulation sites in the control experiment

	Left dPMC			Left pIFG		
	x	*y*	*z*	*x*	*y*	*z*
A	−48	−6	54	−54	4	16
B	−50	−4	54	−58	0	8
C	−46	−6	50	−50	12	16
D	−48	−4	56	−48	8	20
E	−48	−6	54	−52	8	20
F	−50	−6	54	−50	2	24
G	−50	−4	56	−52	8	20
H	−48	2	58	−58	16	14
I	−52	−10	56	−62	8	20
Average ± SD	−48.9 ± 1.8	−4.9 ± 3.2	54.7 ± 2.2	−53.8 ± 4.6	7.3 ± 4.9	17.6 ± 4.7
	Left aSTS					
	x	*y*	*z*			
A	−62	−8	6	—	—	—
B	−50	−4	0	—	—	—
C	−64	−10	2	—	—	—
D	−58	−6	8	—	—	—
E	−64	−12	2	—	—	—
F	−66	−12	4	—	—	—
G	−64	−4	0	—	—	—
H	−62	−6	−6	—	—	—
I	−64	−10	−4	—	—	—
Average ± SD	−61.6 ± 4.9	−8.0 ± 3.2	1.3 ± 4.5	—	—	—

hoc Fisher's PLSD test was performed to compare post-cTBS with pre-cTBS MEP amplitudes at the single post-cTBS time points.

Experiment 3: cTBS effects on behavioral measures and their relationship to speech-related facilitation of articulatory motor cortex excitability. The consequences of cTBS-induced virtual lesions in the left dorsal stream (pSTS, SPT, and pIFG) on speech-processing performance and their relationship to articulatory motor cortex excitability measures were studied in 19 subjects in Experiment 3 (individual stimulation sites, see Fig. 3*A*; MNI coordinates of individual stimulation sites, Table 2), three of which had already participated in Experiment 1. In addition to the cTBS target sites of the left dorsal and ventral auditory stream, the left MOG was targeted as a control site to investigate the topographical specificity of cTBS effects. The MOG was selected as a control site because MOG is involved in visual processing but is not part of the auditory speech-processing network (Murakami et al., 2012; Restle et al., 2012).

Subjects sat in front of a PC screen (Samsung SyncMaster 214T 21.3 inch LCD TFT monitor) and listened to auditory stimuli that were delivered via a headphone. We recorded the speech material from a German female speaker using a digital voice recorder and embedded in white noise because M1 involvement in speech perception is more relevant in noisy conditions (Meister et al., 2007). Signal-to-noise ratios (SNRs) were adjusted for the different tasks using data from a pilot experiment to avoid behavioral ceiling effects. The SNRs were calculated using the following formula:

$$\mathrm{SNR} = 10\log_{10}(\mathrm{S/N})$$

Two listening-selection tests and three listening-repetition tests were conducted. Responses of the listening-selection tests were recorded via button press with the right index finger. Overtly articulated responses during the listening-repetition tests were recorded with a digital voice recorder. Subjects were instructed to try to respond as quickly as possible.

1. *Word-picture matching.* Participants listened to a German noun and were asked to identify their percept by choosing a visual illustration of the object of four possible choices: taken from the standard German speech and language test battery of the NAT publisher that fulfills the criteria published by Snodgrass and Vanderwart (1980) that were presented 5 s after the auditory word. In addition to the target object, a phonological (mostly vowel replacements, e.g., "Maus" vs "Mais") and a semantic distractor (e.g., "Maus" vs "Frosch") and an unrelated picture were visually presented on a four-panel figure. Nouns in the different categories did not differ in word frequency ($p > 0.5$). The matching test was performed at three very low SNRs: 6, ±0, −6 dB. Each concrete, high-frequency, highly imaginable monosyllabic or bisyllabic noun was presented 1 s after starting noise, and 10 different words were presented at each noise level in randomized order, separated by 5 s intertrial intervals.

2.1. *Syllable repetition.* Participants listened to one of six possible syllables (/Ba/, /Da/, /Ga/, /Pa/, /Ta/, /Ka/) in three noise conditions (SNRs: 9, 6, 3 dB) and repeated immediately the syllable they heard. In each trial, one syllable was presented 1 s after starting noise. Each syllable was presented twice per noise level, resulting in 3 noise levels × 6 syllables × 2 repetitions = 36 trials that were delivered in randomized order and separated by 5 s baseline.

2.2. *Syllable identification.* Same conditions as in 2.1, but instead of overtly repeating their percept, participants were instructed to press one of six predefined buttons representing the six possible choices.

3. *Pseudoword repetition.* This task aimed at studying phonological working memory effects given that the phonological load was higher compared with the syllable condition: Participants listened to pseudowords consisting of three syllables each in three noise conditions (SNRs: 12, 9, 6 dB) and immediately repeated them. Each pseudoword was presented for 1 s after starting noise, and 15 different pseudowords were presented at each noise level interleaved with 5 s baseline per trial. Pseudowords and noise levels were randomized. Trials were judged as error trials when one of the repeated syllables was wrong.

4. *Sentence repetition.* Subjects listened to one of 30 German declarative sentences in three noise conditions (SNRs: 9, 6, 3 dB) and immediately repeated them. Each sentence was presented 1 s after starting noise. Noise levels were randomized. Every missed or wrongly repeated word was counted as an error.

For each task, error rates (ERs) and reaction times (RTs) were calculated at each noise level. The word-picture matching task allowed for classification of errors as phonological, semantic, or unrelated. Errors in all other tasks were classified as place of articulation errors (i.e., mistaking the articulatory zone, e.g., a /Ba/ for a /Da/ or a /Pa/ for a /Ta/) or voicing errors (i.e., mistaking the co-occurrence of phonation with articulation of the initial stop consonant, e.g., a /Ba/ for a /Pa/ or a /Da/ for a /Ta/). Perception of place of articulation of initial stop consonants relies mostly on analyses of spectral cues during the initial 20 ms of the acoustic signal (Miller and Nicely, 1955; Blumstein and Stevens, 1979), whereas voicing analysis requires longer segments that carry temporal information (Rosen, 1992; Shannon et al., 1995).

The sessions started with a pre-cTBS recording of orbicularis oris MEPs during listening to speech or control noise and were followed by the cTBS protocol. The five behavioral assessments and post-cTBS MEP recordings during speech and noise perception were performed within 30 min. The procedures of coil navigation were the same as in Experiment 1. The order of cTBS sites was chosen randomly, and consecutive sessions in a given participant were at least 1 week apart.

ER and RT data from word-picture matching were analyzed by a three-way repeated-measures ANOVA with stimulus site (4 levels: pSTS, SPT, pIFG, MOG), noise level (3 levels: low, medium, high), and error type (5 levels: overall, phonological, semantic, unrelated, no response) as within-subject factors. Data from syllable repetition and identification were analyzed by a three-way repeated-measures ANOVA with type of output (2 levels: repetition vs identification), stimulus site (4 levels), and noise level (3 levels) as within-subject factors. Those of pseudoword and sentence repetition were examined using two two-way repeated-measures ANOVAs with stimulus site and noise level as within-subject factors. Absolute MEP amplitude was analyzed by a three-way repeated-measures ANOVA with stimulus site (4 levels), experimental condition (2 levels: listening to sentences vs noise), and time (2 levels: post-cTBS vs pre-TBS) as within-subject factors.

In case of a significant interaction or significant main effect, *post hoc* paired *t* tests were performed using Fisher's PLSD test. We additionally tested for correlations between ER and RT in the five speech perception

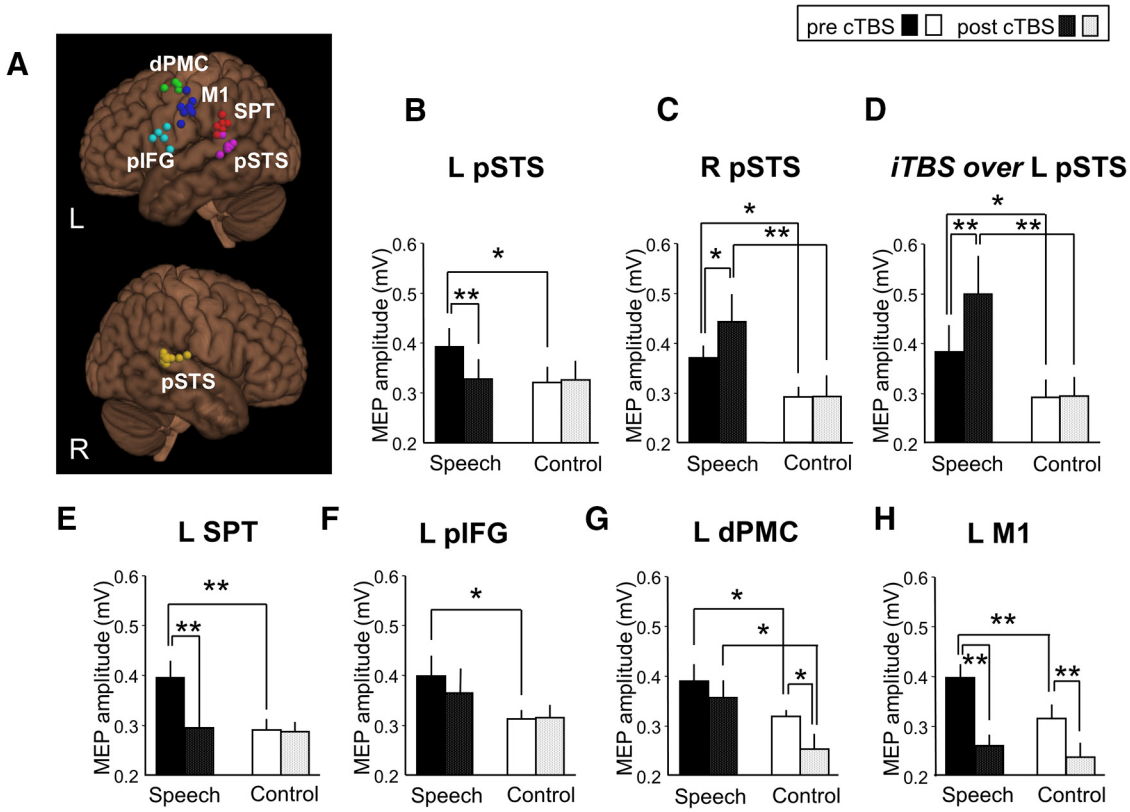

Figure 1. cTBS effects on speech-related facilitation of left articulatory M1 excitability as measured by MEP amplitude (in mV) in the right orbicularis oris muscle. ***A***, Individual stimulation sites over left dPMC (green), pIFG (cyan), M1 (blue), SPT (red), and pSTS (pink) and right pSTS (yellow) are overlaid on an MNI standard brain. Not all stimulation sites are visible due to overlap. ***B***, The speech-related MEP increase (speech: passive listening to German sentences) was reduced after cTBS of the left pSTS with no MEP change in the control condition (control: listening to white noise). cTBS of the right pSTS (***C***) and facilitatory iTBS of the left pSTS (***D***) enhanced speech-related MEP facilitation without any MEP modulation in the control condition. ***E***, cTBS of the left SPT disrupted the speech-related MEP facilitation. ***F***, cTBS of the left pIFG had no significant effect on speech-related MEP facilitation. ***G***, cTBS of the left dPMC had no significant effect on speech-related MEP facilitation but decreased articulatory M1 excitability in the control condition. ***H***, MEPs in speech and control conditions were strongly reduced by cTBS of left articulatory M1. $^{*}p < 0.05$. $^{**}p < 0.01$. Error bars indicate SEM. L, Left; R, right.

tasks with the task-dependent change in MEP amplitude when listening to sentences compared with noise after cTBS of the control region MOG using Pearson's correlation analyses (two-tailed). We applied a Bonferroni correction for multiple comparisons. In case we found a significant correlation, we examined whether this relationship was affected by cTBS of dorsal stream components.

Experiment 4: dorsal stream specificity and frontal parallel pathways in the dorsal stream. Experiment 4 tested 9 subjects who had already participated in Experiment 3. On two occasions, separated by 1 week, participants received cTBS either over the left aSTG or over both the left pIFG and dPMC together (double virtual lesion, individual stimulation sites, see Fig. 6*A*; MNI coordinates of individual stimulation sites, Table 3). aSTG was chosen because it is a constituent of the ventral speech stream (Rauschecker and Scott, 2009). The double-knock-out approach of pIFG/dPMC was chosen because individual cTBS-induced virtual lesioning did not result in significant disruption of speech-related MEP facilitation in articulatory M1 (compare results of Experiment 1). The order of these two conditions was randomized. In the double virtual lesion condition, dPMC and pIFG were targeted immediately one after the other by cTBS, with the order of targets randomized across participants. Parameters and duration of subsequent dPMC and pIFG stimulation were kept identical with the single virtual lesions to avoid effect reversals at short durations of cTBS (Gamboa et al., 2010). This doubles the total amount of overall brain tissue and time the brain is stimulated but keeps the direct effects on the target regions constant.

The behavioral measures obtained during word-picture matching, repetition of syllables, and pseudowords after the virtual lesions of the pIFG/dPMC were compared with those after lesioning aSTG, and also with those after cTBS of the control region MOG from Experiment 3. ERs of word-picture matching were analyzed using a three-way mixed ANOVA with noise level (3 levels) and error type (5 levels) as within-subject factors and stimulus site (3 levels: dPMC/pIFG, aSTG, MOG) as between-subject factor. ERs of syllable and pseudoword repetition were examined using two-way mixed ANOVAs with noise level as within-subject factors and stimulus site as between-subject factor. In addition, we compared behavioral effects of double knock-out of dPMC and pIFG in Experiment 4 versus single knock-out of pIFG in Experiment 3 using the appropriate ANOVAs. MEP amplitudes during listening to speech or noise were analyzed using a three-way mixed repeated-measures ANOVA with experimental condition (2 levels: listening to sentences vs noise) and time (2 levels) as within-subject factors and stimulus site (3 levels: dPMC/pIFG, aSTG, MOG) as between-subject factor. In addition, MEP effects of the double knock-out of dPMC and pIFG in Experiment 4 were compared with the single cTBS over dPMC and pIFG from Experiment 1 in a mixed ANOVA with experimental condition (2 levels: listening to speech vs listening to white noise) as within-subject factor and stimulus site (3 levels: dPMC/pIFG, dPMC, pIFG) as between-subject factor. *Post hoc* paired *t* tests were run in case of a significant interaction or significant main effect using Fisher's PLSD test.

For all of the above statistical analyses, significance was assumed if $p < 0.05$. SPSS for Windows, version 17.0 (SPSS) was used for all statistical analyses.

Results

Experiment 1: effects of repetitive TMS on articulatory motor cortex excitability

Before cTBS, MEPs were facilitated by passive listening to sentences compared with listening to white noise on all occasions (all $p < 0.05$; Fig. 1*B–H*).

The repeated-measures ANOVA showed a significant three-way interaction between experimental condition (2 levels: listening to speech vs white noise), stimulus site (6 levels: left pSTS, right pSTS, left SPT, left pIFG, left dPMC, left M1), and time (2 levels: pre- vs post-cTBS) ($F_{(5,35)} = 5.202$, $p = 0.001$). cTBS over the left pSTS abolished the MEP facilitation induced by listening to speech (differences in MEP amplitude before vs after cTBS: $T = 4.972$, $p = 0.002$) without MEP changes in the control condition ($T = -0.454$, $p = 0.664$; Fig. 1*B*). In contrast, a virtual lesion of the right pSTS enhanced the speech-induced MEP facilitation ($T = -2.508$, $p = 0.040$) without any MEP modulation in the control condition ($T = -0.057$, $p = 0.956$; Fig. 1*C*). This suggests that the right STS physiologically inhibits facilitatory processing in the left dorsal stream in a context-dependent manner. To prove that rTMS could not just interfere with, but also gate speech-induced modulation of articulatory M1 excitability, iTBS, which facilitates stimulated cortex (Huang et al., 2005), was applied to the left pSTS. The repeated-measures ANOVA with main effects of intervention (2 levels: iTBS vs cTBS), experimental condition (2 levels: listening to speech vs white noise), and time (2 levels: pre- vs post-TBS) demonstrated a significant three-way interaction ($F_{(1,7)} = 48.898$, $p < 0.001$). iTBS enhanced the speech-related facilitatory effect on MEP amplitude ($T = -4.960$, $p = 0.002$) without any MEP amplitude change in the control condition ($T = -0.176$, $p = 0.865$; Fig. 1*D*).

cTBS of the left SPT abolished similar to virtual lesioning of pSTS the MEP facilitation induced by listening to speech ($T = 5.136$, $p = 0.001$) without any effect in the control condition ($T = 0.297$, $p = 0.775$; Fig. 1*E*). Although these manipulations proved that integrity of both left pSTS and SPT is necessary for modulation of articulatory M1 excitability during speech perception and, thus, supports the idea that these regions constitute the posterior input into the dorsal speech processing stream, they do not reveal the frontal targets that mediate these effects to M1. Therefore, we studied the consequences of cTBS over the two proposed frontal target regions, the left dPMC and the left pIFG. A virtual lesion of either the dPMC or the pIFG did not block the speech perception related MEP facilitation effect (as in the posterior target regions). Instead, we observed only a nonsignificant reduction in the MEP facilitation ($T = 1.592$, $p = 0.155$ for pIFG, Fig. 1*F*; $T = 0.747$, $p = 0.479$ for dPMC, Fig. 1*G*).

The MEP facilitation induced by listening to speech could only be abolished by virtually lesioning both the dPMC and the pIFG as demonstrated in Experiment 4 ($T = 4.640$, $p = 0.002$; see Fig. 7*A*): triple-interaction of repeated-measures ANOVA with main effects of experimental condition (2 levels: listening to sentences vs noise), stimulus site (3 levels: dPMC/pIFG, aSTG, MOG), and time (2 levels: pre- vs post-cTBS) significant at $F_{(2,34)} = 3.509$, $p = 0.041$. A significant interaction between condition and stimulus site was found when comparing double and single knock-out of dPMC and pIFG ($F_{(2,22)} = 4.203$, $p = 0.028$), which resulted from less MEP facilitation in the double compared with the single knock-out of dPMC and pIFG ($p = 0.028$ and 0.047, respectively). cTBS to the ventral stream region aSTG did not affect speech-related facilitation of motor cortex excitability ($p > 0.2$; see Fig. 7*A*). Of note, single virtual lesions of the other temporal target regions (pSTS and SPT) abolished the speech-related MEP facilitation (see above), which demonstrates the regional specificity of cTBS and argues against a mere quantitative effect of overall brain stimulation dose in the double knock-out condition. This suggests that MEP facilitation by listening to speech can be mediated by either one of the two frontal regions and strongly suggests that these regions operate as two parallel frontal target regions of the left dorsal auditory stream (Hickok and Poeppel, 2007; Mashal et al., 2012).

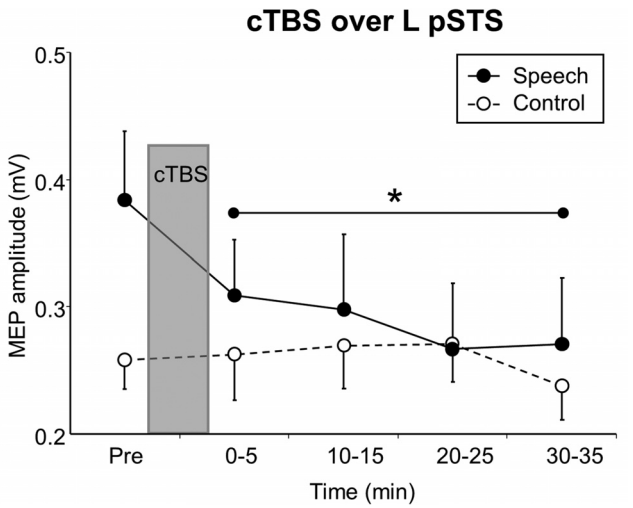

Figure 2. Long-lasting effects of cTBS of left pSTS. cTBS of left pSTS reduced speech-related MEP facilitation for at least 30 min (speech: passive listening to German sentences) without any MEP modulation in the control condition (control: listening to white noise). *$p < 0.05$, Speech-induced MEP amplitudes at all post-cTBS time points are different from those at pre-cTBS (paired *t* test). Error bars indicate SEM.

To control for the efficacy of our virtual lesion, we applied cTBS over left articulatory M1, which significantly reduced MEP amplitude in both speech ($T = 5.721$, $p = 0.001$) and control condition ($T = 4.252$, $p = 0.004$) as expected (Fig. 1*H*). cTBS of the left dPMC, but not cTBS of the left pIFG, also decreased MEP amplitudes in the control condition ($T = 2.612$, $p = 0.035$; Fig. 1*F*,*G*), suggesting direct connectivity between the dPMC and the articulatory M1.

Experiment 2: temporal evolution and duration of cTBS effects

Significant speech-related MEP facilitation was found pre-cTBS only and was disrupted for at least 30 min after cTBS of the left pSTS, whereas cTBS did not modify MEP amplitudes in the control condition. The two-way repeated-measures ANOVA with 2 levels of experimental condition and 5 levels of time revealed a significant interaction ($F_{(4,16)} = 3.868$, $p = 0.022$; Fig. 2). *Post hoc t* tests showed that MEP amplitudes in the listening to speech condition were different at all post-cTBS time points from those at pre-cTBS (Fig. 2, asterisk). The demonstration that cTBS interfered with brain function in a ROI for at least 30 min was a prerequisite for investigating cTBS offline effects on speech-processing performance (see next section).

Experiment 3: cTBS effects on behavioral measures and their relationship to speech-related facilitation of articulatory motor cortex excitability

As expected, behavioral results depended on noise in all regions and conditions, including the control region MOG. For auditory word comprehension (word-picture matching), a three-way repeated-measures ANOVA with stimulus site (4 levels: pIFG, pSTS, SPT, MOG), noise level (3 levels: low, medium, high), and error type (5 levels: overall, phonological, semantic, unrelated, no response) as within-subject factors revealed significant effects of error type ($F_{(4,72)} = 51.708$, $p < 0.001$) and of noise level ($F_{(2,36)} = 81.857$, $p < 0.001$). The effect of error type was explained by much larger phonological and overall errors (Fig. 3*B*,*E*) than the other errors (Fig. 3*C*,*D*,*F*) and justified subsequent separate two-

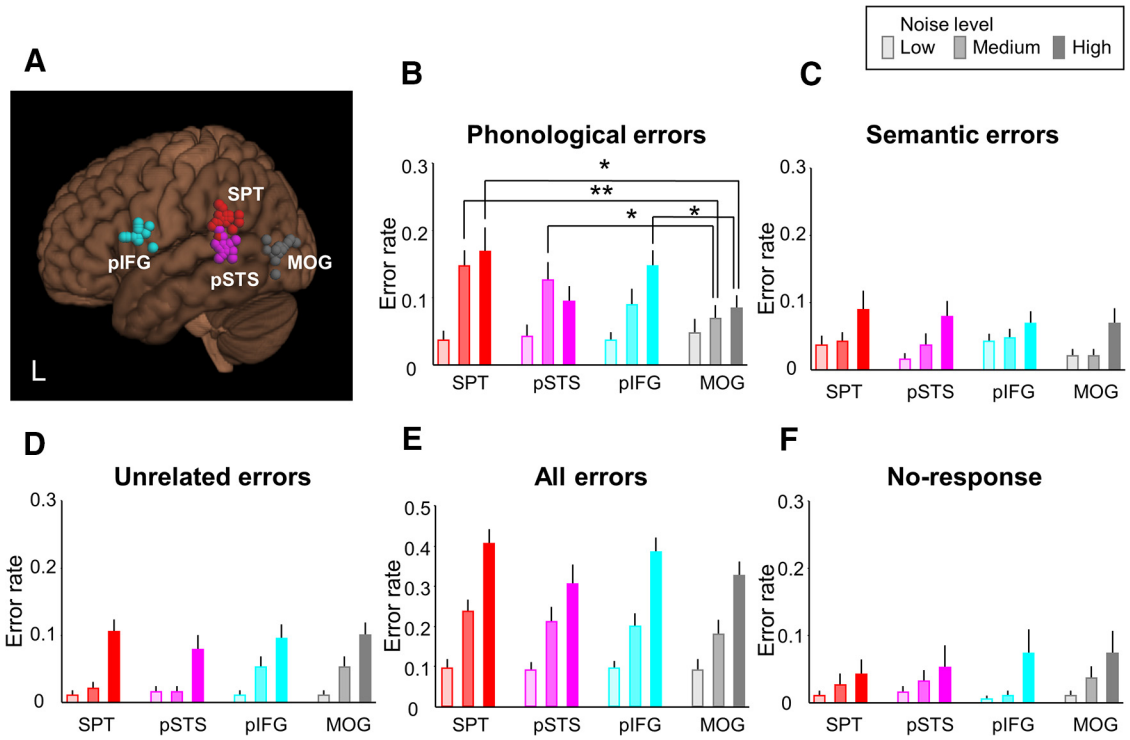

Figure 3. Dorsal stream contribution to auditory word comprehension in noise. ***A***, Individual stimulation sites over left pIFG (cyan), SPT (red), pSTS (pink), and MOG (gray) are overlaid on an MNI standard brain. pIFG, SPT, and pSTS are nodes of the dorsal speech-processing stream, whereas MOG is a control site. Not all stimulation sites are visible due to overlap. L, Left. ***B***, cTBS of SPT and pSTS increased phonological errors for medium-level noise, and cTBS of SPT and pIFG increased phonological errors for high-level noise compared with cTBS of MOG. Semantic (***C***), unrelated errors (***D***), all errors (***E***), and no response (***F***) showed no significant difference between cTBS sites. *$p < 0.05$. **$p < 0.01$. Error bars indicate SEM.

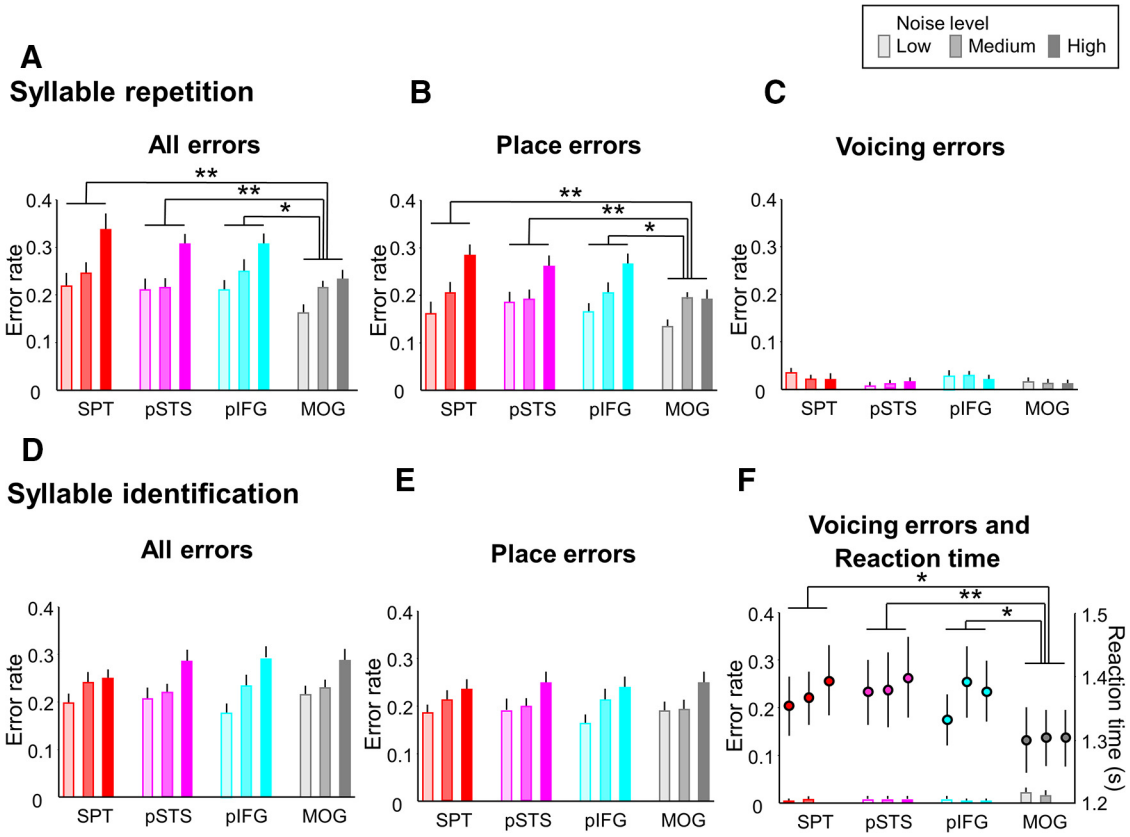

Figure 4. Behavioral cTBS effects on syllable repetition and identification. ***A***, cTBS of left SPT, pSTS, and pIFG increased all errors significantly compared with cTBS of MOG by affecting analysis of place of articulation (***B***) but not analysis of voicing of initial stop consonants (***C***). ***D–F***, No difference between cTBS sites was found for error rates during syllable identification. ***F***, Right *y*-axis, Reaction times for correct trials of syllable identification were longer after cTBS over SPT, pSTS, or pIFG compared with cTBS over MOG. *$p < 0.05$. **$p < 0.01$. Error bars indicate SEM.

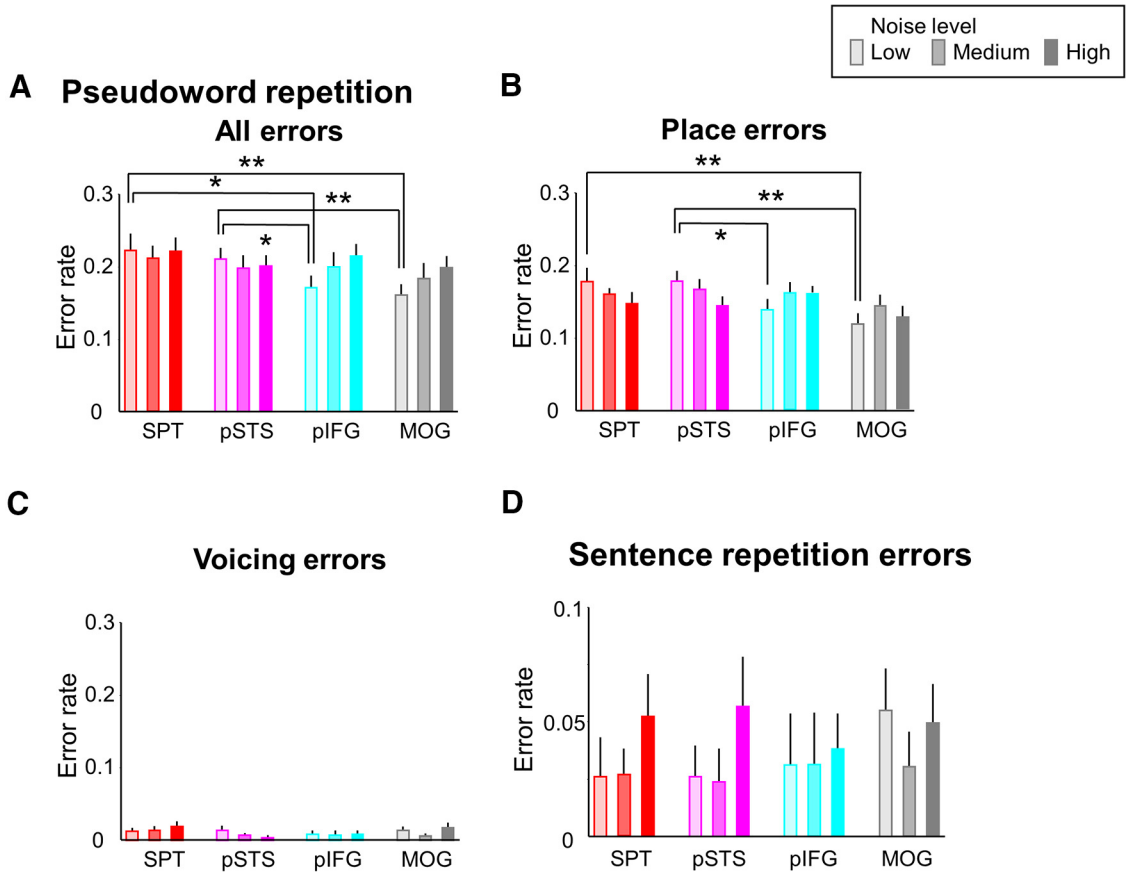

Figure 5. Behavioral cTBS effects on pseudoword and sentence repetition. ***A***, For pseudoword repetition, cTBS of SPT or pSTS increased error rates significantly compared with cTBS of pIFG and MOG at low-level noise. ***B***, ***C***, Nearly all errors were place of articulation and not voicing errors. ***D***, There was no error rate difference between cTBS sites for sentence repetition. *$p < 0.05$. **$p < 0.01$. Error bars indicate SEM.

way repeated-measures ANOVAs divided by error type. A two-way repeated-measures ANOVA with phonological ER as dependent variable and stimulus site (4 levels) and noise level (3 levels) as within-subject factors showed a significant interaction ($F_{(6,108)} = 2.304$, $p = 0.039$). *Post hoc t* test revealed that cTBS of SPT and pSTS increased phonological errors at intermediate noise ($T = 3.525$, $p = 0.002$ for SPT; $T = 2.157$, $p = 0.045$ for pSTS) and that cTBS of SPT and pIFG increased phonological errors at high noise ($T = 2.109$, $p = 0.049$ for SPT; $T = 2.118$, $p = 0.048$ for pIFG) compared with cTBS of MOG (Fig. 3*B*). The other error types (semantic, unrelated errors, all errors, and no response) showed no significant cTBS effect (all $F < 1.8$, $p > 0.1$; Fig. 3*C–F*). No effects on RT were observed (all $F < 1.1$, $p > 0.3$).

For ER during syllable repetition and identification at much higher SNRs compared with the word-picture matching paradigm, a three-way repeated-measures ANOVA with type of output (2 levels), stimulus site (4 levels), and noise level (3 levels) as within-subject factors demonstrated a significant interaction between type of output and stimulus site ($F_{(3,54)} = 6.709$, $p = 0.001$). *Post hoc* testing revealed that ER during syllable repetition after cTBS of SPT, pSTS, and pIFG was larger than those after cTBS over MOG (SPT vs MOG, $T = 3.737$, $p < 0.001$; pSTS vs MOG, $T = 2.689$, $p = 0.009$; pIFG vs MOG, $T = 3.546$, $p = 0.001$; Fig. 4*A*), whereas ER during syllable identification showed no significant difference between sites of cTBS (all $p > 0.3$; Fig. 4*D–F*). Nearly all errors during syllable repetition were place of articulation errors (Fig. 4*B*). RT data showed a significant interaction between type of output and stimulus site in a three-way repeated-measures ANOVA ($F_{(3,54)} = 3.933$, $p = 0.013$). *Post hoc* testing revealed that cTBS in ROIs prolonged RT during syllable identification compared with cTBS of the control region MOG (Fig. 4*F*; SPT vs MOG $T = 2.292$, $p = 0.026$, pSTS vs MOG $T = 3.246$, $p = 0.002$, pIFG vs MOG $T = 2.614$, $p = 0.011$), suggesting that participants with virtual lesions in the dorsal stream processed the acoustic information longer to keep ER comparable with the control condition. No such effect was observed for syllable repetition, which suggests that the visual presentation of the choices in the identification task allowed for additional computation leading to more correct answers.

For ER in pseudoword repetition, a significant stimulus site and noise level interaction was demonstrated ($F_{(6,108)} = 2.751$, $p = 0.016$). *Post hoc* testing revealed that ER increased after cTBS of SPT and pSTS compared with those after cTBS of MOG ($T = 3.584$, $p = 0.002$ for SPT, $T = 3.455$, $p = 0.003$ for pSTS) and of pIFG ($T = 2.642$, $p = 0.017$ for SPT, $T = 2.516$, $p = 0.022$ for pSTS) at low-level noise (Fig. 5*A*). Again, nearly all errors were place of articulation errors (Fig. 5*B*). No effect on RT was observed. cTBS of ROIs did not interfere with sentence repetition (all $F < 1.6$, $p > 0.1$, Fig. 5*D*).

Experiment 4: dorsal stream specificity and frontal parallel pathways in the dorsal stream

For word-picture matching, a three-way mixed ANOVA revealed a significant interaction between error type and stimulus site ($F_{(8,136)} = 3.629$, $p = 0.001$; Fig. 6*B*). *Post hoc* testing revealed that

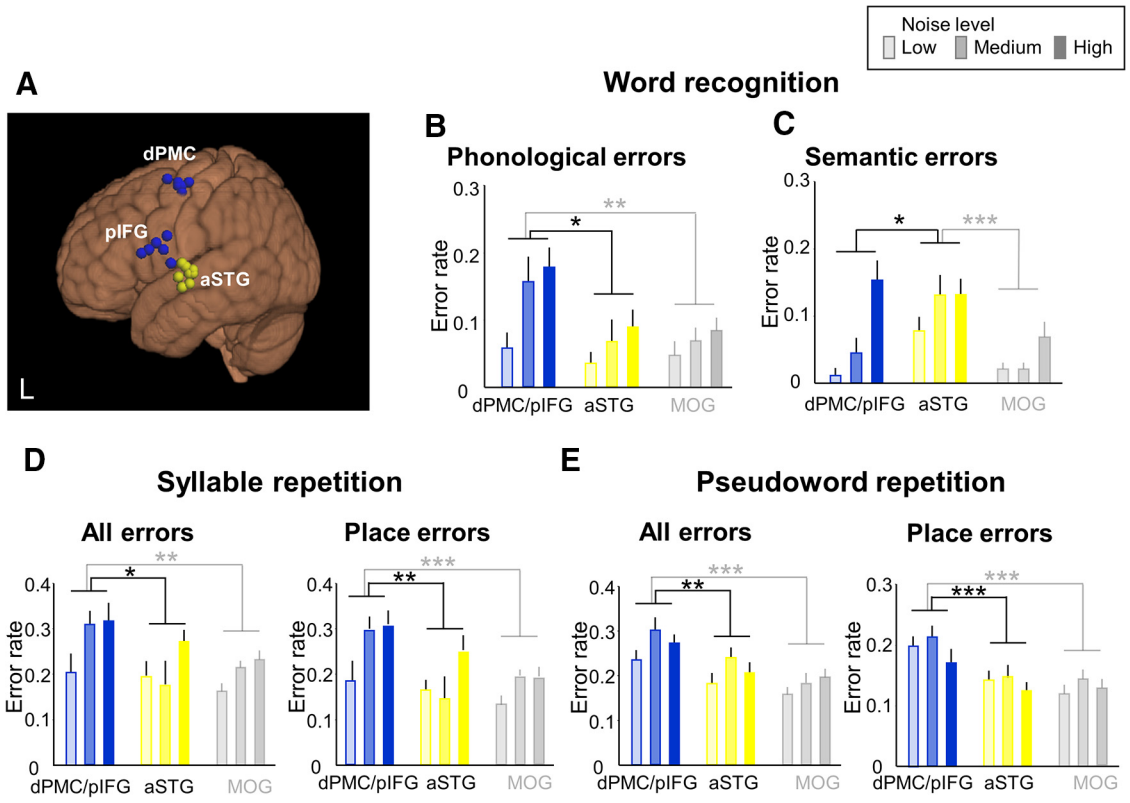

Figure 6. Dorsal stream specificity. ***A***, Individual stimulation sites over left pIFG and dPMC (blue) and aSTG (yellow) are overlaid on an MNI standard brain. aSTG is a control site belonging to the ventral speech-processing stream. Not all stimulation sites are visible due to overlap. ***B***, ***C***, Combined virtual lesions induced by cTBS of the dPMC and the pIFG interfered with phonological but not semantic processing, whereas aSTS/STG lesions increased semantic but not phonological errors during word recognition. ***D***, ***E***, Increased error rates during syllable and pseudoword repetition are induced by the dPMC/pIFG double knock-out but not by virtual lesion of aSTG. The comparison with the control region MOG refers to data of the previous experiment depicted in Figures 3, 4, and 5. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$. Error bars indicate SEM.

a double virtual lesion of both dPMC and pIFG increased phonological errors compared with disruption of aSTG and MOG ($p = 0.017$, $p = 0.005$). In contrast, cTBS of aSTG increased only semantic errors compared with disruption of dPMC/pIFG and MOG ($p = 0.037$, $p < 0.001$; Fig. 6*C*). The other error types showed no significant cTBS effect ($p > 0.08$). For syllable and pseudoword repetition, two-way mixed ANOVAs demonstrated a significant effects of stimulus site ($F_{(2,34)} = 3.473$, $p = 0.042$ for syllable; $F_{(2,34)} = 9.703$, $p < 0.001$ for pseudoword). As expected, double virtual lesions of both dPMC and pIFG increased phonological errors compared with disruption of aSTG and MOG ($p = 0.015$, $p = 0.002$ for syllable; $p = 0.001$, $p < 0.001$ for pseudoword; Fig. 6*D*,*E*) but also compared with single pIFG lesion ($p = 0.112$ for syllable; $p = 0.003$ for pseudoword repetition). Nearly all errors during syllable repetition were place of articulation errors (Fig. 6*D*,*E*).

Single lesions of the aSTG produced more semantic errors than the double knock-out of the dPMC and pIFG, which demonstrates the regional specificity of the cTBS effect and argues against an unspecific effect of overall brain stimulation dose.

Relationship of behavioral measures with task-dependent changes of articulatory motor cortex excitability

In the 19 subjects of Experiment 3, we replicated the finding from the first experiment of a significant interaction between stimulus site (4 levels: pSTS, SPT, pIFG, MOG), experimental condition (2 levels: listening to sentences vs white noise), and time (2 levels: pre-cTBS vs post-cTBS) on speech-related facilitation of MEP amplitude in the right orbicularis oris muscle ($F_{(3,54)} = 2.977$, $p = 0.039$). *Post hoc* testing confirmed that cTBS of SPT and pSTS eliminated facilitatory effects of speech processing on articulatory M1 excitability (MEP amplitude during speech listening pre- vs post-cTBS: $T = 4.972$, $p < 0.001$ for SPT, $T = 3.417$, $p = 0.003$ for pSTS), whereas cTBS over pIFG again produced an intermediate reduction of speech-induced MEP facilitation (MEP amplitude during speech listening pre- vs post-cTBS: $T = 2.437$, $p = 0.025$; Fig. 7*B*). cTBS over the control region MOG did not affect modulation of articulatory M1 excitability during speech perception (pre- vs post-cTBS: $T = -0.038$, $p = 0.970$; Fig. 7*B*). After cTBS over MOG, speech-facilitated MEPs were larger than those after cTBS of the ROIs ($T = 3.986$, $p = 0.001$ for SPT, $T = 2.966$, $p = 0.008$ for pSTS, $T = 3.012$, $p = 0.007$ for pIFG).

There was a significant relationship between speech-induced MEP facilitation (as measured by the MEP amplitude change induced by passive listening to sentences compared with listening to white noise) and behavioral measures of syllable and pseudoword repetition in noise. Those participants who facilitated articulatory M1 excitability more strongly during passive listening to sentences took longer to repeat syllables correctly ($r = 0.579$, $p = 0.009$; Fig. 7*C*, left) and were more correct in repeating pseudowords ($r = -0.585$, $p = 0.009$; Fig. 7*D*, left) after cTBS of the control region MOG. This correlation between MEP change and behavioral measures after virtually lesioning the control region suggests a physiological relationship between modulation of articulatory M1 excitability and phonological processing because TBS in the control region should not affect dorsal stream processing. Importantly, these relationships between electrophysiological and behavioral measures were disrupted upon virtual lesions of the pSTS, SPT, or pIFG (all $p > 0.05$; Fig. 7*C*,*D*, right panels).

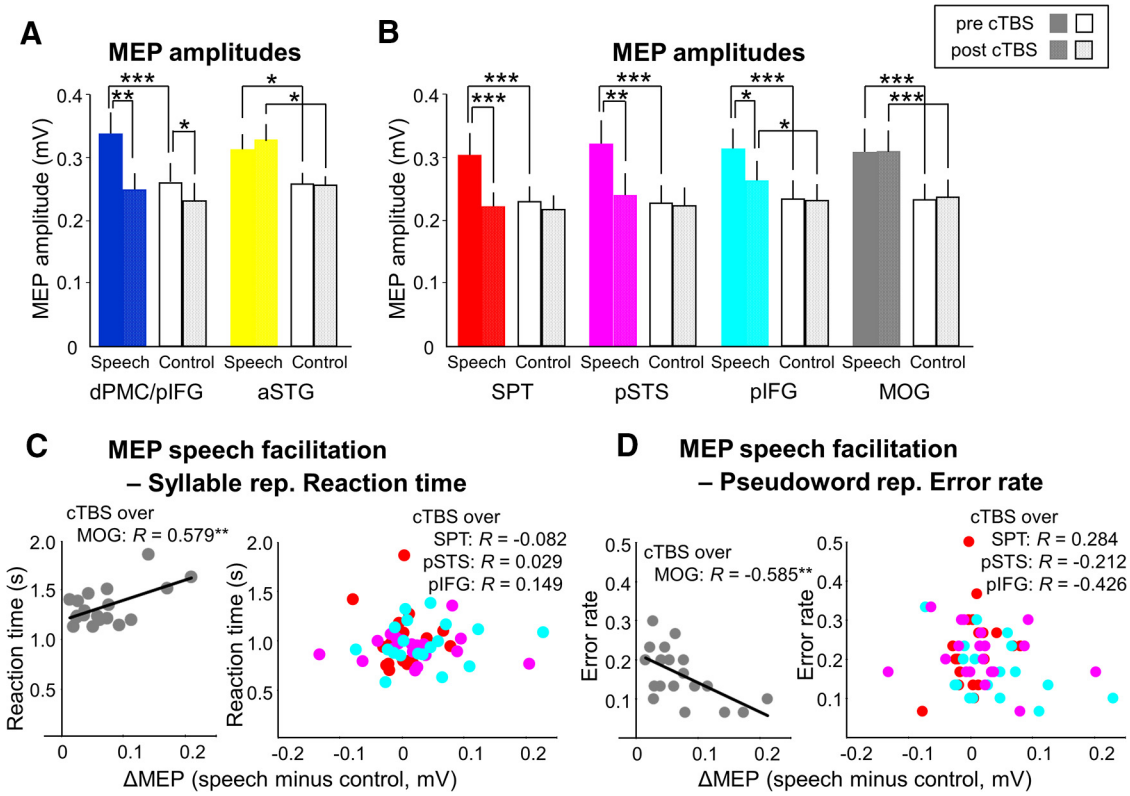

Figure 7. cTBS effects on speech-related modulation of articulatory M1 excitability and correlation analyses with behavioral performance. ***A***, In nine control subjects, the double knock-out of dPMC and pIFG disrupted speech-related MEP facilitation, whereas cTBS of aSTG in the ventral stream did not (speech: passive listening to German sentences; control: listening to white noise). ***B***, In 19 participants, the speech-related MEP facilitation was disrupted after cTBS of the left SPT, pSTS, and pIFG, but not after cTBS of MOG. ***C***, The speech-related MEP facilitation after cTBS of MOG correlated positively with reaction time during correct syllable repetition, whereas cTBS of SPT, pSTS, and pIFG disrupted this relationship. ***D***, The speech-related MEP facilitation after cTBS of MOG correlated negatively with error rates of pseudoword repetition, whereas cTBS of SPT, pSTS, and pIFG decorrelated the electrophysiological measure from behavior. $*p < 0.05$. $**p < 0.01$. $***p < 0.001$. Error bars indicate SEM.

There was no significant correlation between speech-induced MEP facilitation and accuracy of repeating sentences (all $p > 0.05$). Together, these data strongly suggest that additional dorsal stream engagement in passive listening relates to better performance in phonological processes of speech repetition.

Discussion

This study addressed left dorsal stream contributions to speech-related facilitation of articulatory M1 excitability and speech-processing performance and investigated the relationship between these electrophysiological and behavioral effects. Our results suggest that the dorsal stream targets the articulatory M1 via two frontal regions in parallel, the pIFG and the dPMC, whereas the pSTS and SPT constitute essential posterior input regions into the dorsal stream. Our data indicate that speech-related facilitation of articulatory M1 excitability is a reliable measure of dorsal, but not ventral, stream engagement. Importantly, disrupting physiological function in single components of the dorsal stream by cTBS did not result in severe speech perception deficits but rather evoked subtle but significant noise-dependent increases in phonological, particularly place of articulation, errors during speech perception and repetition.

Functional anatomy of the dorsal speech-processing stream

Speech processing models proposed parallel processing in two speech processing streams that connect temporal speech regions to frontal cortices (Hickok and Poeppel, 2007; Rauschecker and Scott, 2009): A ventral stream that originates in auditory cortex and connects anterior temporal lobe regions to anterior parts of Broca's area is thought to mediate transformations between phonology and semantics. A dorsal stream is thought to map sensory speech representations in the posterior superior temporal lobe to articulatory motor cortices in ventral premotor cortex/posterior Broca's area via a sensorimotor interface in area SPT. In our study, articulatory M1 involvement during passive listening to speech was dependent on integrity of the posterior input regions of the dorsal stream, namely, the pSTS and SPT (Hickok and Poeppel, 2007; Murakami et al., 2012), but not the aSTG in the ventral stream. Thus, facilitation of left articulatory M1 excitability during passive listening to speech was not mediated via the ventral stream or the right hemisphere, at least not to an extent to compensate for virtual lesions of the left dorsal stream. We cannot exclude that ventral stream output could also modulate M1 excitability under different conditions, such as linguistic or speech production tasks (see, e.g., Rauschecker, 2011; Hickok, 2012). Although we investigated dorsal stream components separately in this study, there is sufficient empirical evidence from effective connectivity studies (e.g., see Murakami et al., 2012) that they form a task-dependent functional network and thus a "speech processing stream." Our double-knock-out versus individual lesioning data support dorsal stream models in which the dorsal stream is divided into a dorsodorsal stream that connects the SPT to the dPMC and a dorsoventral stream that connects the SPT to the pIFG (Hickok and Poeppel, 2007; Giraud and Poep-

pel, 2012). Such parallel branches within the dorsal stream could also explain why the dPMC usually is a "negative" language site (Sanai et al., 2008) that often can be safely removed during neurosurgery without gross adverse effects on language performance. Nevertheless, lesions of dPMC could affect subtler functions, such as synchronizing motor routines to external sensory cues (Neef et al., 2011), a function that relies upon efficient sensorimotor integration. Thus, the dPMC may be more important in feedback-controlled conditions that require highly precise sensorimotor timing, such as learning how to speak during development or how to utter unknown speech in adulthood (Hartwigsen et al., 2013).

The drastic effects that virtual lesions of dorsal stream components had on speech-related MEP facilitation in articulatory M1 suggest that this electrophysiological measure is ideally suited to study functional connectivity of the left dorsal stream. This was even true for interhemispheric functional connectivity of the pSTS: a virtual lesion in the right pSTS enhanced MEP facilitation while listening to sentences but not when listening to noise. This extends previous findings of symmetrical inhibitory interhemispheric interactions between the pSTS during auditory word recognition to the sentence level (Andoh and Paus, 2011). Facilitatory iTBS of left pSTS produced the same enhancing effect as inhibitory cTBS of right pSTS on speech-related facilitation of articulatory M1 excitability. This specificity of the TBS protocol excludes the possibility that TBS merely interferes with unspecific attentional effects elicited by listening to sentences compared with noise. Together, facilitation and virtual lesions of left hemispheric dorsal speech stream components and of right posterior STS produce pronounced specific electrophysiological effects in left articulatory M1.

On the behavioral level, the same virtual lesions did not abolish the capacity to comprehend or repeat speech. This argues strongly against the motor theory of speech perception (Liberman and Mattingly, 1985). Nevertheless, phonological processing was significantly affected in noisy conditions, and this was found not only during repetition in which sensorimotor mapping is required but also during word and syllable recognition. In case sensory evidence is insufficient to unambiguously link an acoustical percept to a known linguistic item, the brain may gather more evidence from other sensory modalities, such as vision (unavailable to the participants in our study, see e.g., Arnal et al., 2009) or from linguistic context (e.g., via ventral stream processing) (Hickok, 2012). Indeed, sentence repetition was nearly unimpaired by noise and was not correlated with MEP facilitation during passive listening to sentences. In conditions in which neither context nor multisensory integration supports perception as in the tasks studied here, the motor system may contribute to speech perception by internally emulating sensory consequences of articulatory gestures (Du et al., 2014). That is, the impoverished sensory stimulus may call upon sensorimotor associations to create a forward prediction based on an internal model that in turn informs sensory processing (Hickok et al., 2011). This does not mean that this mechanism is necessary to understand speech. It rather implies that consulting internal sensorimotor models ameliorates speech perception when sensory information alone is ambiguous (Meister et al., 2007; Möttönen and Watkins, 2009). Our finding of a significant correlation between phonological behavioral scores and MEP facilitation by speech that is disrupted after virtual lesions of dorsal stream regions (compare Fig. 7*C*,*D*) adds empirical evidence to this notion. Whether such inner emulation happens consciously as in subvocal rehearsal or as part of an automated process is not known. Whatever the exact mechanism, it may be less important in everyday speech perception when contextual information is available.

Noise affects particularly the perception of consonants and reduces primarily the ability to discriminate the place of articulation (Miller and Nicely, 1955). We confirm this finding by showing that place of articulation analysis is affected already at much lower signal-to-noise ratios than vowel perception, and we propose that the left dorsal stream contributes primarily in identifying place of articulation of perceived consonants. This could be explained by the somatotopy of articulator representations, meaning that sensorimotor loops are easily dissociable if different articulators (lips, tongue, palate) are involved. In addition, the motor cortex (together with the basal ganglia and cerebellum) could support phonemic segmentation because it specialized in precise millisecond timing (Kimura, 1993). Importantly, such segmentation requires very fast processing of finely chunked acoustic speech segments, whereas voicing analysis requires bigger chunks carrying temporal information (Rosen, 1992). Our data suggest that the temporal processing underlying voicing analysis relates to the bilateral ventral stream, whereas rapid sensorimotor mapping via the left dorsal stream (Camalier et al., 2012) contributes to perception of fast spectral changes.

One important finding in speech perception neuroimaging studies is the inconsistency of motor activations found during speech perception. A recent study revealed that a substantial interindividual variability in motor engagement in speech perception often diminishes statistical power and thus results in subthreshold motor activations (Szenkovits et al., 2012). More importantly, the authors found that the degree of motor involvement during speech listening as measured by BOLD correlated with offline behavioral phonological working memory scores (pseudoword repetition scores). Phonological working memory has been linked to reverberating sensorimotor loops in the dorsal speech stream (Hickok and Poeppel, 2007; Buchsbaum et al., 2011; Herman et al., 2013). This suggests that those subjects who engage their dorsal stream more strongly in sensorimotor processing during speech perception succeed better in phonological speech tasks or vice versa. Our finding of a correlation of speech-induced MEP facilitation with pseudoword repetition scores and syllable repetition reaction times with electrophysiological data supports this view. Given that we studied only MEP facilitation during passive listening to sentences, we cannot claim that the sublexical tasks studied on the behavioral level would reveal similar relationships to electrophysiology, but we expect this relationship to be even stronger (Möttönen and Watkins, 2009; Herman et al., 2013). This question will need to be addressed in future studies.

In conclusion, we define the left posterior superior temporal region as bottleneck through which articulatory M1 excitability is modulated by speech perception. Our results provide evidence for two parallel branches of a left-lateralized dorsal speech stream, both arising at the posterior end of the Sylvian fissure. Orbicularis oris MEPs can serve as reliable marker of dorsal stream involvement in speech processing. Although sensorimotor mapping via the dorsal stream is not necessary for speech perception, it contributes to phonological analyses in case sensory information is insufficient. It is tempting to speculate that fast processing within the left dorsal stream allows for optimally representing fast changes in the auditory signal and thereby biases speech processing to the left hemisphere.

References

Andoh J, Paus T (2011) Combining functional neuroimaging with off-line brain stimulation: modulation of task-related activity in language areas. J Cogn Neurosci 23:349–361. CrossRef Medline

Arnal LH, Morillon B, Kell CA, Giraud AL (2009) Dual neural routing of visual facilitation in speech processing. J Neurosci 29:13445–13453. CrossRef Medline

Blumstein SE, Stevens KN (1979) Acoustic invariance in speech production: evidence from measurements of the spectral characteristics of stop consonants. J Acoust Soc Am 66:1001–1017. Medline

Buchsbaum BR, Baldo J, Okada K, Berman KF, Dronkers N, D'Esposito M, Hickok G (2011) Conduction aphasia, sensory-motor integration, and phonological short-term memory: an aggregate analysis of lesion and fMRI data. Brain Lang 119:119–128. CrossRef Medline

Camalier CR, D'Angelo WR, Sterbing-D'Angelo SJ, de la Mothe LA, Hackett TA (2012) Neural latencies across auditory cortex of macaque support a dorsal stream supramodal timing advantage in primates. Proc Natl Acad Sci U S A 109:18168–18173. CrossRef Medline

Du Y, Buchsbaum BR, Grady CL, Alain C (2014) Noise differentially impacts phoneme representations in the auditory and speech motor systems. Proc Natl Acad Sci U S A 111:7126–7131. CrossRef Medline

Fadiga L, Craighero L (2003) New insights on sensorimotor integration: from hand action to speech perception. Brain Cogn 53:514–524. CrossRef Medline

Fadiga L, Craighero L, Buccino G, Rizzolatti G (2002) Speech listening specifically modulates the excitability of tongue muscles: a TMS study. Eur J Neurosci 15:399–402. CrossRef Medline

Gamboa OL, Antal A, Moliadze V, Paulus W (2010) Simply longer is not better: reversal of theta burst after-effect with prolonged stimulation. Exp Brain Res 204:181–187. CrossRef Medline

Giraud AL, Poeppel D (2012) Cortical oscillations and speech processing: emerging computational principles and operations. Nat Neurosci 15: 511–517. CrossRef Medline

Groppa S, Oliviero A, Eisen A, Quartarone A, Cohen LG, Mall V, Kaelin-Lang A, Mima T, Rossi S, Thickbroom GW, Rossini PM, Ziemann U, Valls-Solé J, Siebner HR (2012) A practical guide to diagnostic transcranial magnetic stimulation: report of an IFCN committee. Clin Neurophysiol 123: 858–882. CrossRef Medline

Hartwigsen G, Saur D, Price CJ, Baumgaertner A, Ulmer S, Siebner HR (2013) Increased facilitatory connectivity from the pre-SMA to the left dorsal premotor cortex during pseudoword repetition. J Cogn Neurosci 25:580–594. CrossRef Medline

Herman AB, Houde JF, Vinogradov S, Nagarajan SS (2013) Parsing the phonological loop: activation timing in the dorsal speech stream determines accuracy in speech reproduction. J Neurosci 33:5439–5453. CrossRef Medline

Hickok G (2009) Eight problems for the mirror neuron theory of action understanding in monkeys and humans. J Cogn Neurosci 21:1229–1243. CrossRef Medline

Hickok G (2012) The cortical organization of speech processing: feedback control and predictive coding the context of a dual-stream model. J Commun Disord 45:393–402. CrossRef Medline

Hickok G, Poeppel D (2007) The cortical organization of speech processing. Nat Rev Neurosci 8:393–402. CrossRef Medline

Hickok G, Houde J, Rong F (2011) Sensorimotor integration in speech processing: computational basis and neural organization. Neuron 69:407–422. CrossRef Medline

Huang YZ, Edwards MJ, Rounis E, Bhatia KP, Rothwell JC (2005) Theta burst stimulation of the human motor cortex. Neuron 45:201–206. Medline

Kelly YT, Webb TW, Meier JD, Arcaro MJ, Graziano MS (2014) Attributing awareness to oneself and to others. Proc Natl Acad Sci U S A 111:5012–5017. CrossRef Medline

Kimura D (1993) Neuromotor mechanisms in human communication. Oxford: Oxford UP.

Koch G, Fernandez Del Olmo M, Cheeran B, Ruge D, Schippling S, Caltagirone C, Rothwell JC (2007) Focal stimulation of the posterior parietal cortex increases the excitability of the ipsilateral motor cortex. J Neurosci 27:6815–6822. CrossRef Medline

Liberman AM, Mattingly IG (1985) The motor theory of speech perception revised. Cognition 21:1–36. CrossRef Medline

Mashal N, Solodkin A, Dick AS, Chen EE, Small SL (2012) A network model of observation and imitation of speech. Front Psychol 3:84. CrossRef Medline

McConnell J, Quinn JG (2004) Complexity factors in visuo-spatial working memory. Memory 12:338–350. CrossRef Medline

Meister IG, Wilson SM, Deblieck C, Wu AD, Iacoboni M (2007) The essential role of premotor cortex in speech perception. Curr Biol 17:1692–1696. Medline

Miller GA, Nicely PE (1955) An analysis of perceptual confusions among some English consonants. J Acoust Soc Am 27:338–352.

Minagawa-Kawai Y, Cristià A, Dupoux E (2011) Cerebral lateralization and early speech acquisition: a developmental scenario. Dev Cogn Neurosci 1:217–232. CrossRef Medline

Möttönen R, Watkins KE (2009) Motor representations of articulators contribute to categorical perception of speech sounds. J Neurosci 29:9819–9825. CrossRef Medline

Murakami T, Restle J, Ziemann U (2011) Observation-execution matching and action inhibition in human primary motor cortex during viewing of speech-related lip movements or listening to speech. Neuropsychologia 49:2045–2054. Medline

Murakami T, Restle J, Ziemann U (2012) Effective connectivity hierarchically links temporoparietal and frontal areas of the auditory dorsal stream with the motor cortex lip area during speech perception. Brain Lang 122:135–141. CrossRef Medline

Neef NE, Jung K, Rothkegel H, Pollok B, von Gudenberg AW, Paulus W, Sommer M (2011) Right-shift for non-speech motor processing in adults who stutter. Cortex 47:945–954. CrossRef Medline

Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9:97–113. Medline

Rauschecker JP (2011) An expanded role for the dorsal auditory pathway in sensorimotor control and integration. Hear Res 271:16–25. CrossRef Medline

Rauschecker JP, Scott SK (2009) Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nat Neurosci 12:718–724. CrossRef Medline

Restle J, Murakami T, Ziemann U (2012) Facilitation of speech repetition accuracy by theta burst stimulation of the left posterior inferior frontal gyrus. Neuropsychologia 50:2026–2031. CrossRef Medline

Rosen S (1992) Temporal information in speech: acoustic, auditory and linguistic aspects. Philos Trans R Soc Lond B Biol Sci 336:367–373. Medline

Sanai N, Mirzadeh Z, Berger MS (2008) Functional outcome after language mapping for glioma resection. N Engl J Med 358:18–27. CrossRef Medline

Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M (1995) Speech recognition with primarily temporal cues. Science 270:303–304. CrossRef Medline

Snodgrass JG, Vanderwart M (1980) A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. Journal of experimental psychology human learning and memory 6:174–215.

Szenkovits G, Peelle JE, Norris D, Davis MH (2012) Individual differences in premotor and motor recruitment during speech perception. Neuropsychologia 50:1380–1392. CrossRef Medline

Watkins K, Paus T (2004) Modulation of motor excitability during speech perception: the role of Broca's area. J Cogn Neurosci 16:978–987. CrossRef Medline

Watkins KE, Strafella AP, Paus T (2003) Seeing and hearing speech excites the motor system involved in speech production. Neuropsychologia 41: 989–994. CrossRef Medline