

# Motor Representations of Articulators Contribute to Categorical Perception of Speech Sounds

Riikka Möttönen<sup>1,2</sup> and Kate E. Watkins<sup>1,2</sup>

<sup>1</sup>Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, United Kingdom, and <sup>2</sup>Centre for Functional Magnetic Resonance Imaging of the Brain, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, United Kingdom

Listening to speech modulates activity in human motor cortex. It is unclear, however, whether the motor cortex has an essential role in speech perception. Here, we aimed to determine whether the motor representations of articulators contribute to categorical perception of speech sounds. Categorization of continuously variable acoustic signals into discrete phonemes is a fundamental feature of speech communication. We used repetitive transcranial magnetic stimulation (rTMS) to temporarily disrupt the lip representation in the left primary motor cortex. This disruption impaired categorical perception of artificial acoustic continua ranging between two speech sounds that differed in place of articulation, in that the vocal tract is opened and closed rapidly either with the lips or the tip of the tongue (/ba/–/da/ and /pa/–/ta/). In contrast, it did not impair categorical perception of continua ranging between speech sounds that do not involve the lips in their articulation (/ka/–/ga/ and /da/–/ga/). Furthermore, an rTMS-induced disruption of the hand representation had no effect on categorical perception of either of the tested continua (/ba/–/da/ and /ka/–/ga/). These findings indicate that motor circuits controlling production of speech sounds also contribute to their perception. Mapping acoustically highly variable speech sounds onto less variable motor representations may facilitate their phonemic categorization and be important for robust speech perception.

## Introduction

Speech communication requires precise categorization of complex acoustic signals and accurate controlling of articulatory gestures. Growing evidence shows that regions of the premotor and primary motor (M1) cortex, involved in speech production, are also active during speech perception, suggesting that speech perception and production rely partly on the same neural mechanisms (Fadiga et al., 2002; Watkins et al., 2003; Wilson et al., 2004; Pulvermüller et al., 2006; Roy et al., 2008). This view was supported by a recent study in which disruption of the premotor cortex with repetitive transcranial magnetic stimulation (rTMS) impaired recognition of syllables in noise (Meister et al., 2007). The functional role of discrete motor representations of articulators in speech perception is, however, undetermined.

When listening to speech, we sort highly variable acoustic signals into discrete phoneme categories allowed by the language we speak. This fundamental aspect of speech perception—known as categorical perception (CP)—is typically studied by using artificial speech sounds ranging between two phonemes (Liberman et al., 1957). Such continua are created by gradually changing one parameter such as the voice-onset time to vary

voicing (e.g., in a /ga/–/ka/ continuum) or the slope of the formant transition to vary the place of articulation (e.g., in a /ba/–/da/ continuum). Typically, when people listen to these sounds they perceive one phoneme or another rather than something in between. The phoneme category boundaries, thus, divide the acoustic continua into qualitatively discrete regions. Two sounds drawn from one side of the boundary are often perceived to be the “same.” In contrast, two sounds drawn from opposite sides of the boundary are often perceived to be “different”; that is, they are discriminated accurately. It has been proposed that CP reflects the categorical nature of speech production (Fry et al., 1962; Liberman et al., 1967; Liberman and Mattingly, 1985). Articulatory gestures are less variable than acoustic speech signals. For example, phoneme /b/ is produced always by closing the vocal tract with lips, but its acoustic features are highly variable reflecting variability in speech rate and phonetic context, among other factors.

We aimed to determine whether CP of speech sounds relies on motor representations of articulators. Specifically, we tested whether disrupting the neural circuits in M1 cortex that control lip movements with rTMS would impair CP of speech sounds that are produced with lip movements. In experiment 1, CP was tested for a place-of-articulation continuum ranging between lip- and tongue-articulated phonemes (/ba/–/da/) and for a voice-onset-time continuum (/ka/–/ga/). We hypothesized that disruption of the lip, but not hand, representation would impair CP of the former continuum only. To further explore the specificity of the observed effect, the CP of two additional place-of-articulation continua was tested in experiment 2 (/pa/–/ta/ and /da/–/ga/). We hypothesized that disruption of the lip representation would impair CP of the continuum between lip- and tongue-

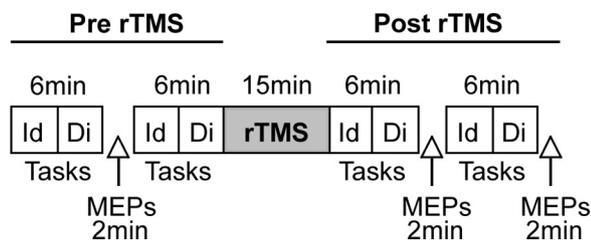
Received Dec. 18, 2008; revised June 8, 2009; accepted July 8, 2009.

This study was supported by the European Commission (Marie Curie Fellowship to R.M.). TMS equipment was funded in part by a grant from the John Fell Oxford University Press Fund to K.W. and Heidi Johansen-Berg and by Medical Research Council funding to K.W. and Wellcome Trust funding to H.J.B. We are grateful to Professor Ruth Campbell for useful discussions at the early stages of this study. We thank Drs. Patricia Gough and Cornelia Stoeckel for help in collecting the data, Iain Wilson for technical assistance, and Drs. Heidi Johansen-Berg and Ingrid Johnsrude for comments on this manuscript.

Correspondence should be addressed to Dr. Riikka Möttönen, Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK. E-mail: riikka.mottonen@psy.ox.ac.uk.

DOI:10.1523/JNEUROSCI.6018-08.2009

Copyright © 2009 Society for Neuroscience 0270-6474/09/299819-07\$15.00/0



**Figure 1.** Experimental design. Participants performed the identification (Id) and discrimination (Di) tasks with two continua before and after a 15 min rTMS train. Stimuli from /ba/–/da/ and /ka/–/ga/ continua were used in experiment 1 and from /pa/–/ta/ and /da/–/ga/ continua in experiment 2. Baseline MEPs were recorded in the middle of the pre-rTMS condition. Post-rTMS MEPs were recorded in the middle and at the end of the post-rTMS condition.

articulated phonemes (/pa/–/ta/), but not CP of the continuum between two tongue-articulated phonemes (/da/–/ga/).

## Materials and Methods

**Participants.** Thirty right-handed native English speakers participated in the study in which either the lip or hand area of left M1 was stimulated. Three participants did not complete the experiment because of discomfort. The data from three participants were excluded because of unreliable motor-evoked potentials (MEPs) or inability to identify stimuli reliably. Data from 24 participants were included in the analyses: experiment 1a ( $n = 8$ , 18–28 years, 4 males), experiment 1b ( $n = 8$ , 18–31 years, 4 males), and experiment 2 ( $n = 8$ , 21–28 years, 3 males). All participants were medication-free and had no personal or family history of seizures or other neurological disorders. Informed consent was obtained from each participant before the experiment. The study was performed under permission from the National Research Ethics Service.

**Procedure.** Lip representations (experiments 1a and 2) and hand representations (experiment 1b) in left M1 cortex were temporarily disrupted by application of a 15 min train of low-frequency rTMS. The success of this inhibition was verified in each participant by comparing the size of MEPs elicited by single pulses of TMS over the representation of the target muscle before and after the rTMS train (Fig. 1). To assess the behavioral effects of the rTMS-induced temporary disruption, participants performed identification and discrimination tasks before and after the rTMS train (Fig. 1). Two acoustic continua were tested in each experiment: /ba/–/da/ and /ka/–/ga/ in experiments 1a and 1b and /pa/–/ta/ and /da/–/ga/ in experiment 2.

**Electromyography recordings.** During TMS of the lip representation, electromyography (EMG) activity from the orbicularis oris muscle was recorded from two surface electrodes (22×30 mm ABRO neonatal electrocardiogram electrodes) attached to the right corners of the lower and upper lip (orbicularis oris muscle). During TMS of the hand representation, EMG activity was recorded from the first dorsal interosseous muscle of the right hand and from the right side of the orbicularis oris muscle. The ground electrode was attached to the forehead in all experiments. The EMG signals were amplified, bandpass filtered (at 0.1–1000 Hz) and sampled (at 5000 Hz) using a CED 1902 amplifier, a CED 1401 analog-to-digital converter, and a PC running Spike software (v. 3; Cambridge Electronic Design). They were stored on the computer for off-line analysis.

To use lower stimulation intensities, cortical excitability was measured in contracted muscles. Participants were trained, therefore, to produce a constant level of contraction of the lip or hand muscles while receiving visual feedback indicating the power of EMG activity. They were trained for ~10 min until a satisfactory constant level of contraction of between 20 and 30% of their maximum voluntary contraction was obtained. Participants were asked to produce this level of contraction while single pulses of TMS were applied over the cortex to determine the active motor threshold and while cortical excitability was measured before and after the rTMS. MEPs were excluded from analysis if the EMG activity in the 100 ms before the TMS pulse indicated outlying (mean  $\pm$  SD) levels of contraction.

**TMS.** All TMS pulses were monophasic, generated by a Magstim 200 and delivered through a 70 mm figure-eight coil connected through a

BiStim module (Magstim). The coil was placed tangential to the skull, such that the induced current flowed from posterior to anterior under the junction of the two wings of the figure-eight coil. The position of the coil over the lateral scalp was adjusted until a robust MEP was observed in the contralateral target muscle.

**rTMS.** Low-frequency (0.6 Hz, subthreshold, 15 min) rTMS was delivered over either the lip or the hand representation of M1 cortex. Fifteen minutes of low-frequency rTMS has been shown to inhibit excitability of M1 cortex (i.e., to reduce MEP amplitudes) for a further 15 min after the end of rTMS (Chen et al., 1997). In each participant, we first determined the active motor threshold: the intensity at which TMS elicited at least 5 of 10 MEPs with an amplitude of at least 200  $\mu$ V when the muscle was contracted at 20–30% of the maximum. The mean active motor threshold (percentage of maximum stimulator output,  $\pm$  SE) for the lip area of left M1 was 54.9% ( $\pm$ 2.6%) in experiment 1a and 51.3% ( $\pm$ 3.0%) in experiment 2. The mean active motor threshold for the hand area of left M1 in experiment 1b was 51.3% ( $\pm$ 2.6%). The intensity of each participant's active motor threshold was used for rTMS. Active motor thresholds are lower than resting motor thresholds. As the rTMS was applied while participants relaxed their muscles the intensities used are subthreshold, that is, they are not sufficiently high enough to elicit MEPs. The EMG signal was monitored throughout to ensure that muscles were relaxed and no MEPs were elicited in the target muscle during rTMS. The coil was replaced after 7.5 min to avoid overheating.

**Single pulse TMS.** To ensure that the rTMS suppressed cortical excitability of M1, we recorded MEPs elicited by single-pulse TMS over the representation of the target muscle before rTMS and again 7 and 15 min after its cessation. At each time point, 20 TMS pulses with random interpulse-intervals of between 5 and 7 s were delivered. The intensity used was determined before the rTMS as that which produced MEPs of mean peak-to-peak amplitude of 1 mV on 10 consecutive trials during muscle contraction. The mean stimulator intensity ( $\pm$ SEM) used to elicit MEPs in the lip muscle was 62.6% ( $\pm$ 2.4%) in experiment 1a and 59.6% ( $\pm$ 3.1%) in experiment 2 and to elicit MEPs in the hand muscle was 57.4% ( $\pm$ 2.4%) in experiment 1b.

**Stimuli.** Four eight-step continua were created using Klatt synthesis: /ba/–/da/, /pa/–/ta/, /ka/–/ga/ and /da/–/ga/ (Klatt, 1980). To create the eight stimuli of the continuum from /ba/ to /da/, the slope of the formant transition was changed by increasing the onset frequency of F2 from 1100 Hz to 1615 Hz and that of F3 from 2250 Hz to 2940 Hz in equal steps. The onset frequency of F1 was 400 Hz in all eight stimuli. The unvoiced continuum from /pa/ to /ta/ was created using the same onset frequencies, but 70 ms of aspiration noise was added to the beginning of the stimuli. To create the eight stimuli of the continuum from /ka/ to /ga/, the voice-onset time was gradually changed by shortening the length of the aspiration noise from 70 to 0 ms in equal steps. The onset frequencies of the formants in all eight stimuli from this continuum were 300 Hz (F1), 1700 Hz (F2), and 1850 Hz (F3). The continuum from /da/ to /ga/ was created by gradually changing the onset frequency of F2 from 1615 Hz to 1700 Hz and that of F3 from 2940 Hz to 1850 Hz in equal steps. The onset frequency of F1 was 400 Hz in all eight stimuli. The length of the formant transition was 50 ms for all stimuli in all continua. During the 280 ms steady-state part, the formant frequencies were 750 Hz (F1), 1200 Hz (F2), and 2500 Hz (F3) for all stimuli.

**Tasks.** In the identification task, all stimuli (from 1 to 8) of a continuum were presented 12 times in a randomized sequence with a stimulus-onset-asynchrony (SOA) of 1500 ms. Participants indicated which syllable they heard after each stimulus by pressing buttons (a two-choice task). In the discrimination task, participants were presented with pairs of stimuli separated by two steps on the continuum (i.e., pairs: 1–3, 2–4, 3–5, 4–6, 5–7, 6–8). Each pair (SOA between sounds 500 ms) was presented 12 times in a randomized sequence with a SOA of 2000 ms. The order of stimuli in each pair was counterbalanced, e.g., pairs 1–3 and 3–1 were both repeated six times. The participants were asked to indicate whether the two syllables sounded the same or different by pressing response buttons. In both tasks, participants pressed buttons with the middle and the index fingers of their left-hand. They were asked to be as accurate and fast as possible. The tasks were controlled with Presentation software (Neurobehavioral Systems). The stimuli were delivered through insert

earphones (Etymotic Research). These also served to protect the participants' hearing during the TMS.

Participants were familiarized with the tasks and the stimuli before the experiment started. All tasks (identification and discrimination with both acoustic continua) were performed before and after the 15 min train of rTMS (Fig. 1). The order of /ka/–/ga/ and /ba/–/da/ tasks in experiment 1 and that of /pa/–/ta/ and /da/–/ga/ tasks in experiment 2 was counterbalanced across subjects; the identification task always preceded the discrimination task and both were completed for one continuum before the identification and discrimination task of the other continuum was tested.

**Analysis of behavioral data.** To estimate CP of the acoustic continua during the identification task, logistic curves were fit to each participant's data to obtain slopes and positions of phonetic category boundaries. After removal of anticipatory responses (reaction time shorter than 200 ms), the proportions of responses were calculated for all eight stimuli in each continuum (for example, proportions of /ba/ responses to each stimulus along /ba/–/da/ continuum). The logistic curves were fit to each participant's pre- and post-rTMS identification data using SPSS software (version 15.0), which uses the following formula:  $E(Y_i) = (1 + \beta_0 \beta_1^i)^{-1}$ . The logarithm of  $\beta_1$  was used as the slope index. The higher the slope index, the steeper the logistic curve (i.e., category boundary). The position of the category boundary was defined as the point along the eight-step continuum corresponding to  $E(0.5)$ .

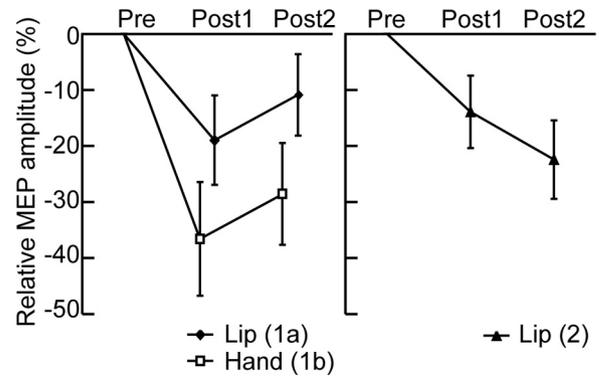
The stimulus pairs presented during the discrimination task were classified as across-category and within-category pairs using the pre-rTMS identification data of each participant. Since the position of the category boundary varied across participants, different pairs were classified as across-category and within-category pairs in different participants. By definition, across-category pairs consist of two stimuli (A and B) that the participant identifies as different syllables. We used the criterion that the difference in response proportions to stimulus A and B had to exceed 0.6 (during pre-rTMS identification task) to be classified as an across-category pair. For example, if the proportion of /ba/ responses to stimulus A was 0.9, then the proportion of /ba/ responses to stimulus B had to be 0.3 or less ( $0.9 - 0.3 = 0.6$ ). One or two stimulus pairs fulfilled this criterion in each participant and were thus classified as across-category pairs. The two pairs furthest away from the category boundary were classified as within-category pairs. Then, the proportions of different responses (given during discrimination tasks) were calculated for these across-category and within-category pairs separately for each individual, to obtain estimates of syllable discrimination in pre- and post-rTMS conditions.

The effects of rTMS-induced disruptions on CP of acoustic continua (slope and position of category boundary and discrimination accuracy of across- and within-category syllables) were statistically tested using ANOVAs. Planned paired *t* tests were performed to compare pre- and post-rTMS conditions (one-tailed as rTMS-induced disruptions were expected to impair CP).

## Results

### rTMS suppressed cortical excitability

Low-frequency rTMS (15 min at 0.6 Hz) successfully suppressed cortical excitability, that is, it temporarily disrupted the lip and hand representations in left M1. This was seen by the reduction of MEP size elicited in either the lip or hand muscles by single pulses of TMS over their respective M1 representations (Fig. 2). In experiment 1a, the mean peak-to-peak amplitude of lip MEPs obtained pre-rTMS was 0.9 mV ( $\pm 0.05$  mV). Seven minutes after the end of the 15 min train of rTMS over the lip area, the mean amplitude was suppressed by 19% ( $t_{(7)} = -2.38, p = 0.025$ ) and this suppression was maintained to the end of the experiment (11% reduction 15 min after rTMS;  $t_{(7)} = -1.49, p = 0.089$ ). In experiment 1b, the mean peak-to-peak amplitude of hand MEPs obtained pre-rTMS was 1.81 mV ( $\pm 0.15$  mV). Seven minutes after the end of the 15 min of rTMS over the hand area, the mean amplitude was suppressed by 37% ( $t_{(7)} = -3.47, p = 0.005$ ) and



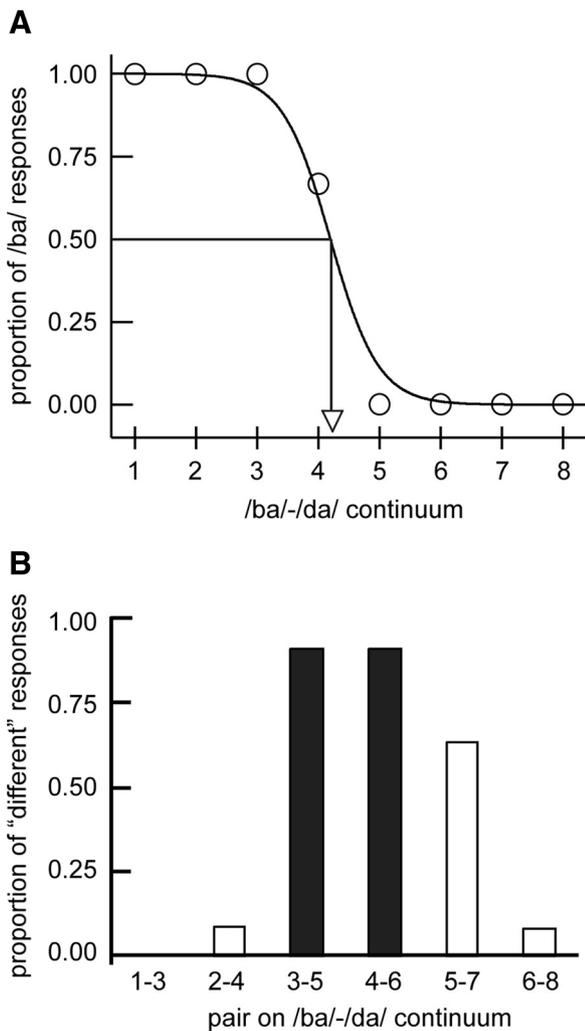
**Figure 2.** Effect of rTMS on cortical excitability. The graph presents mean ( $\pm$ SE) peak-to-peak amplitudes of post-rTMS MEPs in relation to pre-rTMS MEPs. In experiments 1a and 2, rTMS was applied over the lip area of left M1 and MEPs were recorded from the lip muscle. In experiment 1b, rTMS was applied over the hand representation and MEPs were recorded from the hand muscle. The post-rTMS MEPs were recorded  $\sim$ 7 min (Post1) and 15 min (Post2) after the end of the 15 min low-frequency rTMS train.

this suppression was maintained to the end of the experiment (29% reduction 15 min after rTMS;  $t_{(7)} = -2.18, p = 0.033$ ). In experiment 2, the mean peak-to-peak amplitude of lip MEPs obtained pre-rTMS was 1.0 mV ( $\pm 0.05$  mV). The MEPs were suppressed by 14% ( $t_{(7)} = -2.32, p = 0.026$ ) 7 min and by 22% ( $t_{(7)} = -3.46, p = 0.005$ ) 15 min after the end of the rTMS train over the lip area. There were no significant differences in baseline EMG-activity (100 ms before the TMS pulse) between pre- and post-rTMS conditions in any experiment.

### Categorical perception

CP of the acoustic /ba/–/da/ continuum is demonstrated in Figure 3A, which shows data from a participant in a pre-rTMS identification task. This participant consistently identified the stimuli from 1 to 3 as /ba/ and the stimuli from 5 to 8 as /da/. The proportion of /ba/ responses to the stimulus 4 was 0.67. A logistic curve was fit to these data points to estimate the boundary between phoneme categories. In this participant, the slope index of the curve was 1.11 and the position of the abrupt boundary between phoneme categories was between acoustic stimuli 4 and 5 (4.2). Group analyses evaluated the effects of rTMS-induced disruptions on the slope and position of category boundaries derived from the logistic curves fit to each individual's data (Table 1, Fig. 3A).

Discrimination performance for stimulus pairs drawn from the acoustic /ba/–/da/ continuum is demonstrated in Figure 3B, which shows data from the participant in a pre-rTMS discrimination task. The proportion of different responses to pairs 3–5 and 4–6 was 0.92, indicating that the participant accurately discriminated the two stimuli. These across-category pairs comprise stimuli that the participant labeled as different phonemes during the identification task (e.g., stimulus 3 was identified as /ba/ and stimulus 5 was identified as /da/, see Fig. 3A). On the other hand, the within-category pairs (e.g., 1–3 and 6–8) comprise stimuli that the participant labeled as same phonemes during the identification task (e.g., both 1 and 3 were identified as /ba/ and both 6 and 8 were identified as /da/) (Fig. 3A). During the discrimination task, the participant reported that these pairs were the same, indicating that the two stimuli were not discriminated. Group analyses evaluated the effects of rTMS-induced disruptions on proportions of different responses to individually defined across- and within-category pairs (Table 2, Fig. 4B).



**Figure 3.** Performance of a participant on CP tasks. **A**, Identification. The graph shows typical performance in a pre-rTMS identification task, demonstrating categorical perception. The circles show proportions of /ba/ responses to each stimulus along the eight-step /ba/-/da/ continuum. A logistic curve was fit to these data points (solid line). The position of the category boundary is at 4.2 along the /ba/-/da/ continuum (see arrow). **B**, Discrimination. The graph shows typical performance in a pre-rTMS discrimination task, in which the participant was presented with two stimuli and asked to indicate whether they were the same or different. The proportions of different responses are plotted. In this participant, the pair of stimuli 3 and 5 and the pair stimuli 4 and 6 were across-category pairs (black bars).

**Table 1. Identification task: Mean slopes and positions of category boundaries in pre- and post-rTMS conditions (±SE)**

	Slopes		Positions	
	Pre	Post	Pre	Post
Lip (experiment 1a)				
/ba/-/da/	<b>0.92 (0.06)</b>	<b>0.80 (0.08)</b>	4.31 (0.28)	4.14 (0.42)
/ka/-/ga/	0.81 (0.07)	0.93 (0.07)	4.31 (0.19)	4.44 (0.13)
Hand (experiment 1b)				
/ba/-/da/	0.87 (0.06)	0.91 (0.05)	4.28 (0.25)	4.30 (0.18)
/ka/-/ga/	0.94 (0.09)	0.86 (0.07)	4.31 (0.04)	4.45 (0.14)
Lip (experiment 2)				
/pa/-/ta/	0.86 (0.07)	0.88 (0.07)	4.84 (0.25)	4.75 (0.15)
/da/-/ga/	0.68 (0.06)	0.78 (0.08)	4.04 (0.39)	3.84 (0.26)

Logistic curves were fit to the identification data to estimate slopes and positions of category boundaries. Bold indicates significant change from pre- to post-rTMS condition ( $p < 0.05$ ).

**Table 2. Discrimination task: Mean proportions of different responses to across- and within-category pairs in pre- and post-rTMS conditions (±SE)**

	Across-category		Within-category	
	Pre	Post	Pre	Post
Lip (experiment 1a)				
/ba/-/da/	<b>0.73 (0.11)</b>	<b>0.58 (0.10)</b>	0.20 (0.08)	0.24 (0.10)
/ka/-/ga/	0.78 (0.11)	0.84 (0.05)	0.08 (0.02)	0.12 (0.03)
Hand (experiment 1b)				
/ba/-/da/	0.71 (0.10)	0.76 (0.08)	0.16 (0.03)	0.10 (0.05)
/ka/-/ga/	0.78 (0.06)	0.71 (0.06)	0.11 (0.04)	0.13 (0.04)
Lip (experiment 2)				
/pa/-/ta/	<b>0.70 (0.07)</b>	<b>0.53 (0.10)</b>	0.12 (0.04)	0.10 (0.03)
/da/-/ga/	0.62 (0.08)	0.65 (0.08)	0.19 (0.06)	0.15 (0.04)

The proportion of different responses provides a measure of discriminability of stimuli. The across-category stimuli were discriminated more accurately than within-category stimuli. Bold indicates significant change from pre- to post-rTMS condition ( $p < 0.01$ ).

**Effect of rTMS-induced disruptions on category boundaries**

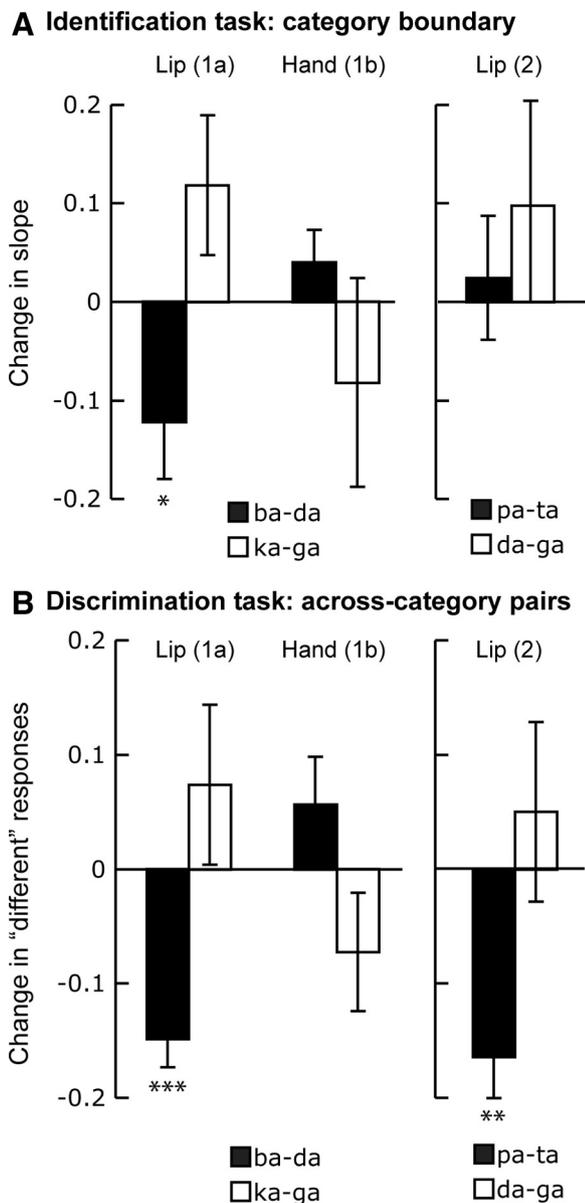
The results of the identification tasks in experiments 1a and 1b can be summarized as follows (Table 1, Fig. 4A): (1) the rTMS-induced disruption of the lip representation in left M1 reduced the slope of the category boundary between /ba/ and /da/; (2) this disruption did not reduce the slope of the boundary between /ka/ and /ga/; (3) the disruption of the hand representation in left M1 had no effect on the slope of either category boundary; (4) the positions of the category boundaries were unaffected by the rTMS-induced disruptions of hand or lip area. The results of the identification tasks in experiment 2 showed that the rTMS-induced disruption of the lip representation did not affect the category boundary for either /pa/-/ta/ or /da/-/ga/ (Table 1, Fig. 4A).

For the slope of the category boundaries in experiments 1a and 1b, ANOVA with stimulation site (lip, hand) as a between-subjects factor and stimulus type (/ba/-/da/, /ka/-/ga/) and rTMS (pre-, post-) as within-subjects factors showed a significant three-way interaction ( $F_{(1,14)} = 9.23, p < 0.01$ ). This was because of a significant two-way interaction between stimulus type and rTMS ( $F_{(1,7)} = 7.04, p < 0.05$ ) for stimulation over the lip area but not the hand area. When the lip representation was disrupted, the slope of the /ba/-/da/ category boundary was significantly shallower (i.e., the slope index was smaller) (pre- vs post-rTMS,  $t_{(7)} = -2.08, p < 0.05$ ), whereas the slope of the /ka/-/ga/ boundary showed a nonsignificant trend to be steeper (i.e., the slope index was larger) (Table 1, Fig. 4A). A three-way ANOVA for the positions of the category boundaries revealed no significant main effects or interactions (Table 1).

A two-way ANOVA for the slopes of the category boundaries obtained in experiment 2 revealed a significant main effect of stimulus type ( $F_{(1,7)} = 6.65, p < 0.05$ ). This was because of a steeper slope of the /pa/-/ta/ than /da/-/ga/ category boundary in this group of participants (means across conditions: /pa/-/ta/  $0.87 \pm 0.05$ , /da/-/ga/  $0.73 \pm 0.05$ ). The main effect of rTMS and the interaction between stimulus type and rTMS were nonsignificant. A two-way ANOVA for positions of category boundaries did not reveal significant main effects or interactions.

**Effect of rTMS-induced disruptions on discrimination of speech sounds**

The results of the discrimination tasks in experiments 1a and 1b can be summarized as follows (Table 2, Fig. 4B): (1) the rTMS-induced disruption of the lip representation in left M1 impaired the discrimination accuracy of across-category sounds from the /ba/-/da/ continuum; (2) this disruption did not impair discrim-



**Figure 4.** Effect of rTMS on categorical perception. **A**, Changes in the slopes of the category boundaries. The graph presents mean ( $\pm$ SE) differences between conditions (post-rTMS–pre-rTMS) in the slope indices derived from the logistic curves fit to each participant's identification data. **B**, Changes in the proportions of across-category pairs perceived as different. The graph presents mean ( $\pm$ SE) differences between conditions (post-rTMS–pre-rTMS) in the proportions of different responses to individually defined across-category pairs. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

ination accuracy for across-category sounds from the /ka/–/ga/ continuum; (3) this disruption did not affect discrimination accuracy of within-category sounds from either continuum; (4) the disruption of the hand representation in left M1 had no effect on discrimination of across- or within-category sounds from either continuum. The results of the discrimination tasks in experiment 2 showed that: (1) the rTMS-induced disruption of the lip representation impaired the discrimination accuracy of across-category sounds from the /pa/–/ta/ continuum, but not from the /da/–/ga/ continuum; (2) this disruption did not affect discrimination of within-category sounds from either continuum (Table 2, Fig. 4B).

For the proportions of different responses to across-category pairs obtained in experiment 1, ANOVA with stimulation site

(lip, hand) as a between-subjects factor and stimulus type (/ba/–/da/, /ka/–/ga/) and rTMS (pre-, post-) as within-subjects factors showed a significant three-way interaction ( $F_{(1,14)} = 15.60$ ,  $p = 0.001$ ). Furthermore, a two-way ANOVA for stimulation over the lip area (i.e., experiment 1a) revealed a significant interaction between stimulus type and rTMS ( $F_{(1,7)} = 10.56$ ,  $p < 0.05$ ). When the lip representation was disrupted, the proportion of across-category /ba/–/da/ pairs perceived as different was significantly decreased (pre- vs post-rTMS,  $t_{(7)} = -5.93$ ,  $p < 0.001$ ), but there was no significant change in the proportion of different responses to the across-category /ka/–/ga/ pairs (Table 2, Fig. 4B). The main effect of stimulus type was also significant ( $F_{(1,7)} = 6.53$ ,  $p < 0.05$ ; means across conditions: /ba/–/da/  $0.66 \pm 0.11$ , /ka/–/ga/  $0.82 \pm 0.08$ ), suggesting that the /ka/–/ga/ stimuli were more easily discriminated than the /ba/–/da/ stimuli in this group of participants. This was primarily because of a significant difference in the post-rTMS condition ( $t_{(7)} = 3.85$ ,  $p < 0.01$ ), whereas a small difference in the pre-rTMS condition was not significant. A two-way ANOVA for stimulation over the hand area (i.e., experiment 1b) revealed no significant main effects or interactions. A three-way ANOVA for the within-category pairs revealed no significant main effects or interactions (Table 2).

A two-way ANOVA for the across-category pairs obtained in experiment 2 showed a marginally significant interaction between stimulus type (/pa/–/ta/, /da/–/ga/) and rTMS (pre-, post-) ( $F_{(1,7)} = 4.52$ ,  $p = 0.07$ ). The disruption of the lip representation significantly decreased the proportion of different responses to across-category /pa/–/ta/ pairs (pre- vs post-rTMS,  $t_{(7)} = -3.33$ ,  $p < 0.01$ ). In contrast, the proportion of different responses to /da/–/ga/ pairs did not change. A two-way ANOVA for the within-category pairs revealed no significant main effects or interactions (Table 2).

## Discussion

We addressed the question of whether motor regions that control movements of articulators play a role in speech perception. The lip representation in left M1 cortex was localized in each participant by recording MEPs from the lip muscle and then targeted with low-frequency rTMS. After rTMS, (1) excitability of the lip representation was disrupted, i.e., MEP sizes were reduced and (2) CP of speech sounds that involve lips in their articulation was impaired. Our findings shed light on the functional role of the motor cortex in speech perception and suggest that mapping highly variable acoustic signals onto discrete motor representations of articulators contributes to speech perception. Such mapping could improve intelligibility of speech when it is compromised for example because of noise.

In experiment 1, an rTMS-induced disruption of the lip representation impaired CP of the acoustic /ba/–/da/ continuum in two ways: the slope of the phoneme category boundary between /b/ and /d/ was shallower and discrimination of these phonemes was harder. Thus, performance changes in both identification and discrimination tasks indicate that during the disruption the acoustic /ba/–/da/ continuum was perceived more continuously, that is, less categorically. This is likely because of impaired detection of the place of articulation of speech sounds; the /b/ and /d/ sounds are produced by closing and opening the vocal tract at the lips and with the tip of the tongue behind the teeth, respectively. This interpretation is supported by the finding that disruption of the lip representation did not impair CP of the /ka/–/ga/ continuum; neither /k/ nor /g/ is produced with the lips.

We expect that disruption of the motor representation of the larynx that is important for controlling voicing would impair CP

of voice-onset-time continua (e.g., /ka/–/ga/). A recent functional magnetic resonance imaging (MRI) study found two activation peaks for laryngeal phonation: a ventromedial peak located deep in the central sulcus and a more superficially located dorsolateral peak (Brown et al., 2008). To examine the role of the larynx representation in speech perception, these areas would first need to be individually localized with functional MRI or by recording MEPs from the laryngeal muscles with needle electrodes and then targeted with TMS. The results of the current study showed a nonsignificant trend toward an enhancement of CP of the voice-onset-time continuum /ka/–/ga/ when the lip representation was disrupted: the slope of the category boundary was steeper and the discrimination of the across-category syllables was more accurate. This might reflect facilitation (or disinhibition) of the larynx representation, adjacent to the inhibited lip representation.

Importantly, an rTMS-induced disruption of the hand representation in left M1 cortex had no effect on CP of either continuum in experiment 1, indicating that CP of the /ba/–/da/ continuum was specifically affected by the disruption of the motor representation of an articulator (i.e., lips). This is in agreement with a previous single-pulse TMS study showing that listening to speech modulates excitability of the motor representation of the lip but not that of the hand (Watkins et al., 2003).

Results of experiment 2 give further support for the view that motor representations of articulators contribute to CP of speech sounds in an articulatory feature-specific manner. Disruption of the lip representation impaired discrimination of /pa/ and /ta/ that are articulated with lip and tongue movements, respectively, replicating the results of experiment 1b. Furthermore, in agreement with our hypothesis, this disruption did not impair identification and discrimination of speech sounds that are articulated with differing tongue movements (i.e., /da/ and /ga/). Contrary to our expectations, we however failed to reduce the slope of the category boundary between unvoiced lip- and tongue-articulated phonemes (i.e., /pa/ and /ta/) in experiment 2, suggesting that discrimination of speech sounds is more sensitive to motor disruptions than identification.

Our findings are in line with an rTMS study showing that disruption of dorsal premotor cortex impaired recognition of syllables with varying articulatory features (/pa/, /ta/, /ka/) in noise (Meister et al., 2007). The present rTMS study complements these findings by showing that (1) disrupting a primary motor representation of an articulator, but not that of hand, affects CP, a fundamental aspect of speech perception and (2) that this effect is sensitive to the articulatory features of speech sounds.

A recent double-pulse TMS study of D'Ausilio et al. (2009) showed that stimulation of the motor representations of the tongue and lip before speech sounds affects their recognition: stimulation of the lip representation facilitated recognition of lip-articulated phonemes (/p/ and /b/), whereas stimulation of the tongue representation facilitated tongue-articulated phonemes (/t/ and /d/). This effect, however, is opposite to the one of the present study (facilitation vs impairment), which could be because of a different type of TMS (online double-pulse TMS vs off-line low-frequency rTMS) (Devlin and Watkins, 2007). Either way, our results together with those of D'Ausilio et al. (2009) demonstrate that TMS can be used to stimulate selectively motor representation of different articulators in M1 and modulate perception of speech sounds in an articulatory feature-specific manner. This demonstrates relatively high spatial accuracy of TMS. It is, however, possible that the behavioral effects are at least partly

caused by TMS-induced modulation in brain regions that are functionally connected with the targeted region in M1 (for example, the premotor cortex).

The critical role of motor representations in speech perception is one of the central claims of the motor theory of speech perception (Liberman et al., 1967; Liberman and Mattingly, 1985). Our findings can be interpreted to support a weak version of this claim. It should be noted that CP was not completely abolished by the motor disruption in the present study. This is not surprising given that rTMS-induced disruptions typically change reaction times or error rates rather than completely abolish behavior (Devlin and Watkins, 2007). Furthermore, current views on speech perception propose that it is mediated by multiple parallel mechanisms using both nonmotor (e.g., acoustic) and motor representations (Scott and Johnsrude, 2003; Davis and Johnsrude, 2007; Hickok and Poeppel, 2007). A complementary role for the motor system in speech perception is supported by the findings of Taylor (1979), showing that although patients with focal lesions of the ventral central sulcus (sensory and motor representations of the face) are impaired at spelling and phoneme recognition relative to patients with temporal or frontal lobe lesions, they still perform at above chance levels. Our findings, together with earlier findings, show that the human motor cortex plays an important role in speech processing. This role is likely to be complementary, but not necessary as suggested by the motor theory of speech perception. Further studies are needed to determine the conditions in which the motor system contributes to speech perception and to clarify its role in the processing of natural speech.

Neuroimaging studies on processing of speech sounds have highlighted the role of temporal and parietal regions. The left posterior superior temporal cortex shows greater responses to speech than to nonspeech sounds (Binder et al., 2000). Furthermore, sine-wave analogues of sounds elicit stronger activity in this region when they are perceived as speech (i.e., phonetically and categorically) than when they are perceived as nonspeech (i.e., nonphonetically and continuously) (Dehaene-Lambertz et al., 2005; Möttönen et al., 2006; Desai et al., 2008). The supramarginal gyrus is also thought to be involved in categorical processing of speech sounds (Raizada and Poldrack, 2007). In humans, both superior temporal cortex and supramarginal gyrus are densely connected with the lateral prefrontal cortex, via white matter tracts (Parker et al., 2005). This dorsal/posterior stream thus provides a link between speech processing mechanisms based on acoustic/phonological representations and motor articulatory representations (Scott and Johnsrude, 2003; Davis and Johnsrude, 2007; Hickok and Poeppel, 2007). Further studies are needed to determine how the temporal and parietal regions interact with frontal motor regions during speech perception.

The idea of tight coupling of perception and production has been under extensive investigation since the discovery of “mirror neurons” in the monkey ventral premotor cortex (di Pellegrino et al., 1992). There is indeed growing evidence that in both monkey and human brain, controlling of one's own hand and mouth actions and coding of seen and heard actions of others are mediated partly by the same neural circuits (Fadiga et al., 1995; Hari et al., 1998; Kohler et al., 2002; Ferrari et al., 2003; Gazzola et al., 2006). These findings have led to a proposal that actions of others are understood by motor simulation, i.e., by mapping them onto the perceiver's own motor representations (Rizzolatti et al., 2001; Gallese, 2003). Such mapping could aid interpersonal communication, including speech communication. Interestingly, human motor and premotor cortex has been shown to map also action-

related words somatotopically, demonstrating semantic somatotopy (for review, see Pulvermüller 2005). The present findings suggest that disruption in somatotopic mapping of speech sounds in motor cortex can modulate their perception, giving further support for the view that motor representations play a role in perception.

## References

- Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN, Possing ET (2000) Human temporal lobe activation by speech and non-speech sounds. *Cereb Cortex* 10:512–528.
- Brown S, Ngan E, Liotti M (2008) A larynx area in the human motor cortex. *Cereb Cortex* 18:837–845.
- Chen R, Classen J, Gerloff C, Celnik P, Wassermann EM, Hallett M, Cohen LG (1997) Depression of motor cortex excitability by low-frequency transcranial magnetic stimulation. *Neurology* 48:1398–1403.
- D'Ausilio A, Pulvermüller F, Salmas P, Bufalari I, Begliomini C, Fadiga L (2009) The motor somatotopy of speech perception. *Curr Biol* 19:381–385.
- Davis MH, Johnsrude IS (2007) Hearing speech sounds: top-down influences on the interface between audition and speech perception. *Hear Res* 229:132–147.
- Dehaene-Lambertz G, Pallier C, Serniclaes W, Sprenger-Charolles L, Jobert A, Dehaene S (2005) Neural correlates of switching from auditory to speech perception. *Neuroimage* 24:21–33.
- Desai R, Liebenthal E, Waldron E, Binder JR (2008) Left posterior temporal regions are sensitive to auditory categorization. *J Cogn Neurosci* 20:1174–1188.
- Devlin JT, Watkins KE (2007) Stimulating language: insights from TMS. *Brain* 130:610–622.
- di Pellegrino G, Fadiga L, Fogassi L, Gallese V, Rizzolatti G (1992) Understanding motor events: a neurophysiological study. *Exp Brain Res* 91:176–180.
- Fadiga L, Fogassi L, Pavesi G, Rizzolatti G (1995) Motor facilitation during action observation: a magnetic stimulation study. *J Neurophysiol* 73:2608–2611.
- 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.
- Ferrari PF, Gallese V, Rizzolatti G, Fogassi L (2003) Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *Eur J Neurosci* 17:1703–1714.
- Fry DB, Abramson AS, Eimas PD, Liberman AM (1962) The identification and discrimination of synthetic vowels. *Lang Speech* 5:171–189.
- Gallese V (2003) The manifold nature of interpersonal relations: the quest for a common mechanism. *Philos Trans R Soc Lond B Biol Sci* 358:517–528.
- Gazzola V, Aziz-Zadeh L, Keysers C (2006) Empathy and the somatotopic auditory mirror system in humans. *Curr Biol* 16:1824–1829.
- Hari R, Forss N, Avikainen S, Kirveskari E, Salenius S, Rizzolatti G (1998) Activation of human primary motor cortex during action observation: a neuromagnetic study. *Proc Natl Acad Sci U S A* 95:15061–15065.
- Hickok G, Poeppel D (2007) The cortical organization of speech processing. *Nat Rev Neurosci* 8:393–402.
- Klatt D (1980) Software for cascade/parallel formant synthesized. *J Acoust Soc Am* 67:971–995.
- Kohler E, Keysers C, Umiltà MA, Fogassi L, Gallese V, Rizzolatti G (2002) Hearing sounds, understanding actions: action representation in mirror neurons. *Science* 297:846–848.
- Liberman AM, Mattingly IG (1985) The motor theory of speech perception revised. *Cognition* 21:1–36.
- Liberman AM, Harris KS, Hoffman HS, Griffith BC (1957) The discrimination of speech sounds within and across phoneme boundaries. *J Exp Psychol* 54:358–368.
- Liberman AM, Cooper FS, Shankweiler DP, Studdert-Kennedy M (1967) Perception of the speech code. *Psychol Rev* 74:431–461.
- 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.
- Möttönen R, Calvert GA, Jääskeläinen IP, Matthews PM, Thesen T, Tuomaine J, Sams M (2006) Perceiving identical sounds as speech or non-speech modulates activity in the left posterior superior temporal sulcus. *Neuroimage* 30:563–569.
- Parker GJ, Luzzi S, Alexander DC, Wheeler-Kingshott CA, Ciccarelli O, Lambon Ralph MA (2005) Lateralization of ventral and dorsal auditory-language pathways in the human brain. *Neuroimage* 24:656–666.
- Pulvermüller F (2005) Brain mechanisms linking language and action. *Nat Rev Neurosci* 6:576–582.
- Pulvermüller F, Huss M, Kherif F, Moscoso del Prado Martin F, Hauk O, Shtyrov Y (2006) Motor cortex maps articulatory features of speech sounds. *Proc Natl Acad Sci U S A* 103:7865–7870.
- Raizada RD, Poldrack RA (2007) Selective amplification of stimulus differences during categorical processing of speech. *Neuron* 56:726–740.
- Rizzolatti G, Fogassi L, Gallese V (2001) Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat Rev Neurosci* 2:661–670.
- Roy AC, Craighero L, Fabbri-Destro M, Fadiga L (2008) Phonological and lexical motor facilitation during speech listening: a transcranial magnetic stimulation study. *J Physiol Paris* 102:101–105.
- Scott SK, Johnsrude IS (2003) The neuroanatomical and functional organization of speech perception. *Trends Neurosci* 26:100–107.
- Taylor L (1979) Psychological assessment of neurosurgical patients. In: *Functional neurosurgery* (Rasmunsses T, Marino R, eds), pp 166–180. New York: Raven.
- Watkins KE, Strafella AP, Paus T (2003) Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia* 41:989–994.
- Wilson SM, Saygin AP, Sereno MI, Iacoboni M (2004) Listening to speech activates motor areas involved in speech production. *Nat Neurosci* 7:701–702.