

Auditory-Motor Processing of Speech Sounds

Riikka Möttönen¹, Rebekah Dutton¹ and Kate E. Watkins^{1,2}

¹Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK and ²Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB), University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, UK

Address correspondence to Riikka Möttönen, Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK. Email: riikka.mottonen@psy.ox.ac.uk

The motor regions that control movements of the articulators activate during listening to speech and contribute to performance in demanding speech recognition and discrimination tasks. Whether the articulatory motor cortex modulates auditory processing of speech sounds is unknown. Here, we aimed to determine whether the articulatory motor cortex affects the auditory mechanisms underlying discrimination of speech sounds in the absence of demanding speech tasks. Using electroencephalography, we recorded responses to changes in sound sequences, while participants watched a silent video. We also disrupted the lip or the hand representation in left motor cortex using transcranial magnetic stimulation. Disruption of the lip representation suppressed responses to changes in speech sounds, but not piano tones. In contrast, disruption of the hand representation had no effect on responses to changes in speech sounds. These findings show that disruptions within, but not outside, the articulatory motor cortex impair automatic auditory discrimination of speech sounds. The findings provide evidence for the importance of auditory-motor processes in efficient neural analysis of speech sounds.

Keywords: auditory cortex, mismatch negativity, motor cortex, sensorimotor, transcranial magnetic stimulation

Introduction

Speech comprehension requires that a listener discriminates accurately between speech sounds belonging to distinct phonological categories. The early stages of speech sound discrimination rely critically on the auditory cortex. Whether the regions in the motor cortex that control movements of the articulators also support processing of speech sounds is under debate (Galantucci et al. 2006; Lotto et al. 2009; Scott et al. 2009; Wilson 2009; Pulvermüller and Fadiga 2010; Hickok et al. 2011). Since the movements of the articulators that modulate the shape of the vocal tract are constrained and less variable than acoustic speech signals, generation of internal motor models of the articulatory gestures of the speaker might aid in categorization of his/her speech sounds (Lieberman et al. 1967; Lieberman and Mattingly 1985; Davis and Johnsrude 2007; Möttönen and Watkins forthcoming).

Growing evidence shows that listening to speech activates the motor cortex (Fadiga et al. 2002; Watkins et al. 2003; Wilson et al. 2004; Pulvermüller et al. 2006). Moreover, recent transcranial magnetic stimulation (TMS) studies provide evidence that the premotor cortex and the representations of articulators (i.e. the lips and tongue) in the left primary motor (M1) cortex contribute to the discrimination of ambiguous

speech sounds (Meister et al. 2007; D'Ausilio et al. 2009; Möttönen and Watkins 2009). In these studies, participants performed demanding speech identification or discrimination tasks, while intelligibility of the speech sounds was compromised by adding noise (Meister et al. 2007; D'Ausilio et al. 2009) or by manipulating formant transitions (Möttönen and Watkins 2009). Although stimulation of the articulatory motor system influenced task performance in these studies, the neural mechanisms underlying these influences remain unknown. For example, it is unknown whether the TMS-induced disruptions in the articulatory motor system modulate auditory processing of speech sounds or whether they affect task-related processes. Also, since behavioral tasks direct the listener's attention to distinctive features of speech sounds, it is unknown whether the articulatory motor cortex influences processing of speech sounds that are outside the focus of attention.

Discrimination of speech and nonspeech sounds can be examined without behavioral tasks by using electroencephalography (EEG) to record responses elicited by occasional changes in sound sequences (Näätänen et al. 2001). These mismatch negativity (MMN) responses are generated in the auditory cortex 100–200 ms after the onset of sound. The MMN responses to phonetic changes in speech sounds are enhanced for phonetic contrasts in the perceiver's native language (Näätänen et al. 1997; Winkler et al. 1999; Diaz et al. 2008). The strongest evidence supporting the independence of MMN responses from attention and task demands comes from the studies that have recorded them in sleeping newborns (Cheour et al. 1998; Cheour et al. 2002), comatose patients (Fischer et al. 1999) and anesthetized rats (Ahmed et al. 2011).

Here, we used EEG and TMS to determine whether the motor cortex contributes to automatic discrimination of speech and nonspeech sounds in the auditory cortex. We stimulated either the lip or hand representation in the left M1 cortex using a 15-min train of low-frequency repetitive TMS, which suppresses motor excitability for up to 20 min after the end of stimulation (Chen et al. 1997; Möttönen and Watkins 2009). Using EEG, we recorded MMN responses to changes in sound sequences, immediately after TMS and in a baseline condition with no TMS. Participants watched a silent film while EEG was recorded and were instructed to ignore the sounds. The combination of TMS and EEG allowed us to determine whether the automatic processes underlying auditory discrimination are independent of the motor cortex or whether they are modulated by TMS-induced disruptions within the motor cortex.

Materials and Methods

Participants

Forty-four right-handed native English speakers participated in this study. Each participant took part in 1 of the 4 experiments (Experiment 1: 10 participants, 24 ± 1.2 years, 8 males; Experiment 2: 12 participants, 27 ± 3.0 years, 7 males; Experiment 3: 12 participants, 26 ± 0.8 years, 9 males; Experiment 4: 10 participants, 24 ± 1.49 years, 5 males). Data from an additional 6 participants were excluded from the analyses because of artifacts in the EEG signal or technical problems during the experiment. 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 prior to the experiment. The study was carried out under permission from the National Research Ethics Service.

Procedure

A 15-min train of low-frequency repetitive TMS was applied over the lip (Experiments 1, 3, and 4) or hand (Experiment 2) representation in the left primary motor cortex. This train reduces excitability of the targeted representation for at least 15 min following the end of the repetitive TMS train (Möttönen and Watkins 2009). To assess the effects of the TMS-induced motor disruption on automatic discrimination of speech sounds, event-related potentials to a 14-min oddball sound sequence were recorded immediately after the TMS train and during a baseline recording. The oddball sequence included 2 infrequent (probability = 0.1 for each) sounds and 1 frequent (probability = 0.8) sound. There were always at least 3 frequent sounds between infrequent sounds. The oddball sequence included 140 repetitions of each infrequent sound. Event-related potentials were also recorded to 2 control sequences. Each control sequence included 400 repetitions of one of the sounds (probability = 1.0) that were used as infrequent sounds in the oddball sequence. The stimulus onset asynchrony was 600 ms in all sequences. The participants watched a silent movie during all recordings in order to focus their attention away from the sound sequence. Half of the experiments started with baseline EEG recordings (oddball and control sequences in a counter-balanced order), which was followed by the TMS train and post-TMS EEG recording (oddball sequence). Half of the experiments started with the TMS train and post-TMS EEG recording (oddball sequence). In these experiments, baseline EEG recordings (oddball and control sequences in a counter-balanced order) were started 50–60 min after the end of the TMS trains when the motor cortex had recovered from the disruption.

Stimuli

In Experiments 1 and 2, the oddball sequence consisted of infrequent “ba” and “ga” and frequent “da” syllables. A female native speaker of British English produced these syllables. The syllables were edited to create 3 stimuli with equal durations (100 ms) and intensity. The steady-state parts of the stimuli corresponding to the vowel sound (i.e. the last 74 ms) were identical in all 3 stimuli. In Experiment 3, the oddball sequence consisted of a frequent “da” stimulus that was identical to the one used in Experiments 1 and 2; the infrequent stimuli differed from it in either duration (170 ms) or intensity (−6 dB). In Experiment 4, the oddball sequence consisted of piano tones (middle C). The frequent tone was edited to have the same duration as the frequent speech stimulus in Experiments 1–3 (100 ms); the infrequent tones differed from it in either duration (170 ms) or intensity (−6 dB) as in Experiment 3. All sounds (except the infrequent sounds with reduced intensity) were played via headphones at ~65 dB SPL.

Transcranial Magnetic Stimulation

Monophasic TMS pulses were generated by 2 Magstim 200s and delivered through a 70-mm figure-8 coil connected through a BiStim module (Magstim, Dyfed, UK). The position of the coil over the lateral scalp was adjusted until a robust motor-evoked potential (MEP) was observed in the contralateral target muscle.

Low-frequency (0.6 Hz, sub-threshold, 15-min) repetitive TMS was delivered over either the lip or the hand representation of the left M1 cortex. In each participant, we determined the active motor threshold (aMT), that is, the minimum intensity at which TMS elicited at least 5 out of 10 MEPs with an amplitude of at least 200 μ V when the target muscle was contracted at 20–30% of the maximum. The electromyography (EMG) signal (passband 1–1000 Hz, sampling rate 5000 Hz) was recorded from the hand (*first dorsal interosseous*) muscle of the right hand and from the right side of the lip (*orbicularis oris*) muscle. The mean aMT (percentage of maximum stimulator output \pm standard error) for the lip area of left M1 was 58.3% ($\pm 2.6\%$) in Experiment 1, 51.8% ($\pm 2.3\%$) in Experiment 3 and 54.4% ($\pm 1.9\%$) in Experiment 4. The mean aMT for the hand area of left M1 in Experiment 2 was 50.5% ($\pm 2.0\%$). The intensity of each participant's aMT was used for repetitive TMS, while the participant was relaxed (i.e. the stimulation was sub-threshold). The stimulation intensity was below 66% for each subject. None of the participants reported experiencing discomfort during the stimulation. During repetitive TMS, the EMG signals were monitored to ensure that muscles were relaxed and no MEPs were elicited in the target muscle. The coil was held manually in a fixed position. It was replaced after 7.5 min to avoid overheating.

Electroencephalography Recordings and Analyses

The EEG (passband 0.1–100 Hz, sampling rate 500 Hz) was recorded with 11 electrodes using an electrode cap (Fz, FCz, Cz, CPz, Pz, F1, F2, P1, P2, left, and right mastoids) in an electrically shielded room. The TMS coil limited the number of electrodes that could be used, since the electrodes could not be placed directly under the coil. The reference electrode was placed on the tip of the nose. Horizontal and vertical eye movements were monitored with electro-oculography recordings. Electrode impedances were kept below 10 k Ω .

The continuous raw data files were re-referenced to the mean of 2 mastoids to improve the signal-to-noise ratio of the MMN responses and digitally high-pass filtered with a cutoff frequency of 1 Hz. Then, 400-ms epochs including a prestimulus baseline of 100 ms were averaged for each stimulus in each condition. The epochs containing amplitude changes exceeding $\pm 70 \mu$ V were removed before averaging. Also the epochs for the first 10 stimuli in each sequence and for the first standard stimulus after each deviant were removed. Finally, the averaged signals were digitally low-pass filtered with a cutoff frequency of 30 Hz.

The baseline (with no TMS) MMN responses were obtained using 2 methods. In Method 1, the responses to frequent sounds were subtracted from responses to the infrequent sounds presented in the same oddball sequence. In Method 2, the responses to stimuli presented frequently in the control sequence (with no TMS) were subtracted from responses to the identical sounds presented infrequently in the oddball sequences. The advantage of Method 1 is that responses to both infrequent and frequent sounds were recorded simultaneously. A possible confound of this method is that the infrequent and frequent sounds differ in their acoustic properties and could therefore elicit different responses. Method 2 compares responses between acoustically identical sounds presented either in a control sequence or as an infrequent sound in the oddball sequence; therefore, these identity MMN responses must reflect discrimination of speech sounds, rather than acoustic differences between them (Jacobsen and Schröger 2001, 2003; Kujala et al. 2007). Method 2 was used, therefore, to confirm that genuine MMN responses were elicited in the baseline conditions. The post-TMS MMN responses were calculated using Method 1 only, because only oddball sequences were presented immediately after TMS. Since the duration of the TMS-induced disruption is short (up to 20 min), there was no time to present the control sequences during the motor disruption.

MMN responses (and identity MMN responses) to all infrequent sounds in all experiments were maximal at the electrode site FCz. This is in agreement with the typical distribution of the response (Näätänen et al. 2001; Näätänen et al. 2011). Therefore, FCz was selected for statistical analyses. The main aim of the analyses was to test whether the TMS-induced disruption modulated MMN responses by comparing MMN responses (calculated using Method 1) between

baseline and post-TMS conditions. First, we ran a paired *t*-test (2-tailed) at each time point 0–300 ms after the onset of sound to compare baseline and post-TMS MMN responses. The aim of these analyses was to determine at which latencies (if any) the MMN responses differed from each other. In order to reduce the likelihood of false-positives due to a large number of *t*-tests, we considered MMNs to be significantly different when the *P*-values were lower than 0.05 at 10 (= 20 ms) or more consecutive time points (Guthrie and Buchwald 1991).

The latency of the peak amplitude of MMN response at FCz was determined for each infrequent sound in each experiment from the mean of baseline and post-TMS grand-average MMN responses. The mean amplitudes of the MMN responses were calculated as the mean voltage across a 40-ms window centered at this peak latency. In order to test whether significant MMN responses were elicited in baseline and post-TMS conditions, we compared MMN mean amplitudes (see above) with zero with *t*-tests. We also calculated grand-averages of baseline identity MMN responses and calculated the mean amplitudes across a 40-ms window centered at the peak latency at FCz. The mean amplitudes were then compared with zero with 1-sample *t*-tests to test whether significant identity MMN responses were elicited in the baseline condition. To assess the effects of TMS-induced disruptions in the lip and hand representation areas in the motor cortex on MMN responses to “ba” and “ga” sounds, ANOVAs with a within-subject factor TMS (baseline vs. post-TMS) and a between subjects factor Experiment (1 vs. 2) were carried out. To assess the effects of TMS-induced in the lip representation of the motor cortex on MMN responses to intensity and duration changes in speech and nonspeech sounds, ANOVAs with a within-subject factor TMS (baseline vs. post-TMS) and a between-subjects factor Experiment (3 vs. 4) were carried out. Pairwise *t*-tests were used in posthoc comparisons.

Results

Experiments 1 and 2: Effects of Disruptions in the Motor Lip and Hand Representations on MMN Responses to Changes in Speech Sounds

First, we examined whether TMS-induced disruptions in the motor system modulate MMN responses to phonetic changes in speech sounds. In Experiment 1, we applied TMS over the lip representation, and in Experiment 2, we applied TMS over the hand representation in the left M1 cortex. In both experiments, we recorded responses to infrequent “ba” and “ga” sounds that were presented among frequent “da” sounds. Both infrequent sounds elicited robust MMN responses at the FCz electrode site, confirming that they were discriminated from the frequent sounds (Fig. 1, Tables 1 and 2). Disruption of the lip representation significantly suppressed the MMN response to “ba” 166–188 ms after the onset of sound and to “ga” 170–210 ms after the onset of sound (Fig. 1). Disruption of the hand representation had no effect on MMN responses to “ba” and “ga” (Fig. 1).

To further test whether the modulations of the MMN responses were sensitive to the location of the TMS-induced disruption in the left M1 cortex, we carried out 2-way ANOVAs for MMN mean amplitudes (Table 1, Fig. 1). For MMN responses to “ba”, ANOVA showed a significant interaction between TMS and Experiment ($F_{1,20} = 4.60, P < 0.05$). The TMS-induced disruption of the lip representation significantly suppressed MMN responses to “ba” ($P < 0.05$), whereas the TMS-induced disruption of the hand representation did not significantly modulate these responses. For MMN responses to “ga”, ANOVA also showed a significant interaction between TMS and Experiment ($F_{1,20} = 7.15, P < 0.05$). The TMS-induced disruption of the lip representation

significantly suppressed MMN responses to “ga” ($P < 0.01$), whereas the TMS-induced disruption of the hand representation did not significantly modulate these responses.

Experiments 3 and 4: Effects of Disruption in the Motor Lip Representation on MMN Responses to Changes in Speech and Non-Speech Sounds

Next, we examined whether disruption of the motor lip representation also modulates MMN responses to acoustic (i.e. non-phonetic) changes in speech sound sequences and whether these modulations are specific to speech. In Experiment 3, we presented “da” syllables and in Experiment 4 we presented piano tones (middle C). In both experiments, sound sequences included infrequent sounds that differed from the frequent ones in either duration or intensity. Both types of acoustic changes in speech and tone sequences

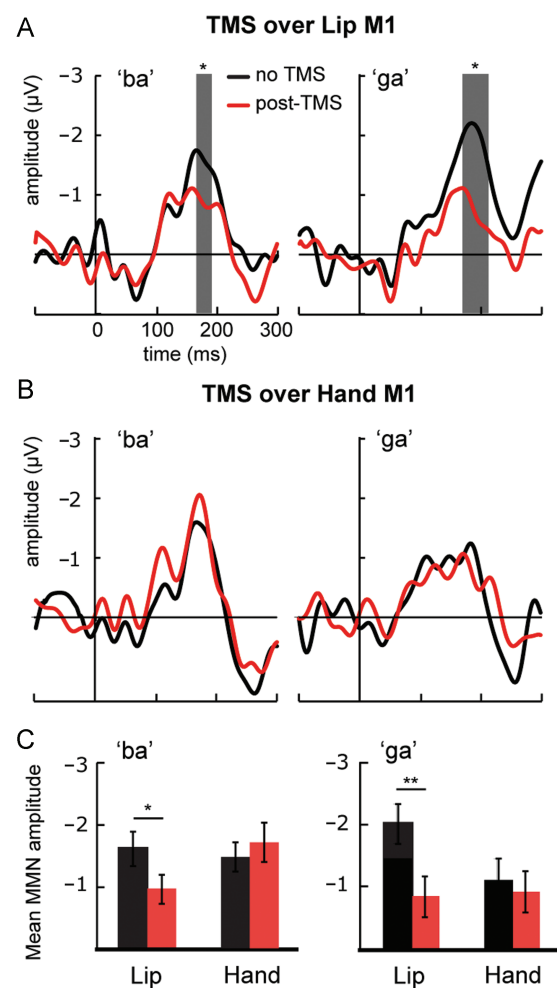


Figure 1. Effect of TMS on MMN responses to phonetic changes in Experiments 1 (Lip) and 2 (Hand). The MMN responses were obtained by subtracting the responses to frequent “da” sounds (probability = 0.8) from the responses to infrequent “ba” (probability = 0.1) and “ga” (probability = 0.1) sounds. Left: Grand-average MMN responses to “ba” at FCz electrode. Right: Grand-average MMN responses to “ga” at FCz electrode. (A) Effect of TMS-induced disruption of the motor lip representation on MMN responses in Experiment 1. (B) Effect of TMS-induced disruption of the motor hand representation on MMN responses in Experiment 2. The gray area indicates the time periods during which the baseline (black) and post-TMS (red) responses differed significantly from each other (sequential *t*-tests). (C) The mean amplitudes (± standard error) of MMN responses (see Table 1). Paired *t*-tests were used in statistical comparisons. * $P < 0.05$, ** $P < 0.01$ (2-tailed).

Table 1MMN peak latencies (ms) and mean amplitudes (μV , $\pm\text{SEM}$) at FCz

	Peak	No TMS		Post-TMS	
		Amplitude	<i>t</i>	Amplitude	<i>t</i>
Experiment 1: TMS over lip M1, <i>n</i> = 10					
"ba"	168	-1.61 (0.28)	-5.79***	-0.96 (0.32)	-3.02**
"ga"	180	-2.06 (0.33)	-6.33***	-0.87 (0.33)	-2.61*
Experiment 2: TMS over hand M1, <i>n</i> = 12					
"ba"	170	-1.47 (0.33)	-4.51***	-1.71 (0.24)	-7.21***
"ga"	176	-1.13 (0.32)	-3.52**	-0.91 (0.33)	-2.75**
Experiment 3: TMS over lip M1, speech, <i>n</i> = 12					
Duration	180	-5.35 (0.71)	-7.50***	-5.14 (0.84)	-6.14***
Intensity	194	-2.33 (0.33)	-7.14***	-1.35 (0.48)	-2.79*
Experiment 4: TMS over lip M1, tones, <i>n</i> = 10					
Duration	194	-5.13 (0.50)	-10.5***	-4.55 (0.49)	-9.23***
Intensity	186	-1.38 (0.36)	-3.82**	-1.69 (0.35)	-4.77***

Note: The MMN mean amplitudes were calculated as the mean voltage across a 40-ms window centred at the peak latency in the grand-average response at FCz. The MMN amplitudes were compared to zero with 2-tailed *t*-tests.

P* < 0.05.*P* < 0.01.****P* < 0.001.**Table 2**Identity MMN peak latencies (ms) and mean amplitudes (μV , $\pm\text{SEM}$) at FCz

	Peak	Amplitude	<i>t</i>
Experiment 1			
"ba"	166	-1.20 (0.29)	-4.12**
"ga"	180	-1.43 (0.32)	-4.46***
Experiment 2			
"ba"	160	-1.16 (0.39)	-3.27**
"ga"	160	-1.33 (0.30)	-4.81***
Experiment 3			
Duration	182	-4.80 (0.82)	-5.85***
Intensity	194	-2.26 (0.41)	-5.46***
Experiment 4			
Duration	198	-4.41 (0.46)	-9.41***
Intensity	180	-1.60 (0.38)	-4.14**

Note: The identity MMN responses were obtained by subtracting the responses to sounds presented in control sequences (probability = 1.0) from the responses to identical infrequent sounds (probability = 0.1) presented in oddball sequences. The mean amplitudes were calculated as the mean voltage across a 40-ms window centred at the peak latency in the grand-average response at FCz. The MMN amplitudes were compared to zero with 2-tailed *t*-tests.

P* < 0.05.*P* < 0.01.****P* < 0.001.

elicited robust MMN responses at the electrode site FCz (Fig. 2, Tables 1 and 2). Disruption of the motor lip representation significantly suppressed MMN to intensity changes in speech sounds 148–170 ms and 182–202 ms after the onset of sound. There were no significant differences between baseline and post-TMS MMN responses to duration changes in speech sounds. MMN responses to intensity and duration changes in tones were unaffected by TMS (Fig. 2).

To further test whether the TMS-induced disruption of the lip representation specifically modulated the MMN responses to changes in sequences of speech sounds, we carried out 2-way ANOVAs for MMN mean amplitudes (Table 1, Fig. 2). For MMN responses to duration changes, ANOVA did not show any significant main effects or interactions. For MMN responses to intensity changes, ANOVA showed a significant interaction between TMS and Experiment ($F_{1,20} = 5.05$,

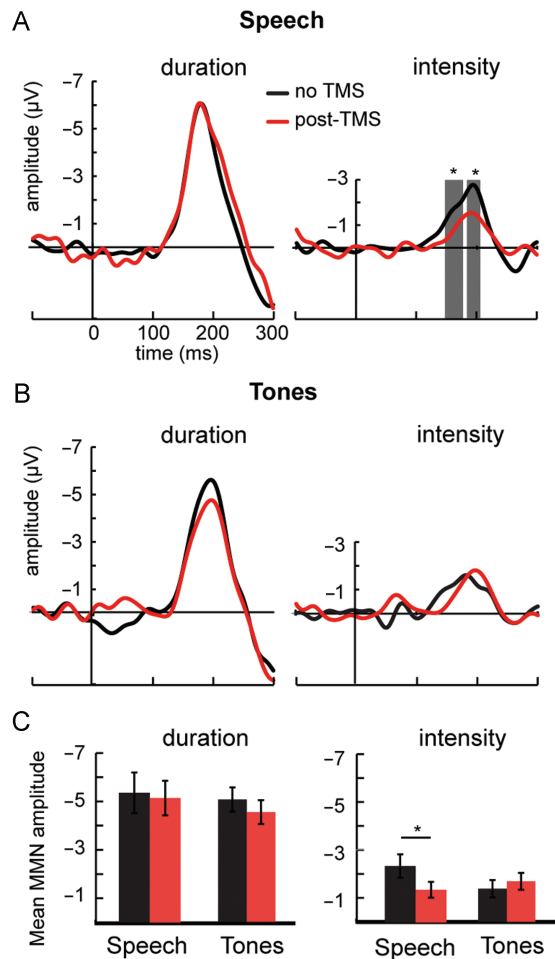


Figure 2. Effect of TMS on MMN responses to duration and intensity changes in Experiments 3 (Speech) and 4 (Tones). The MMN responses were obtained by subtracting the responses to frequent sounds (probability = 0.8) from the responses to infrequent sounds that differed in duration (+70 ms, probability = 0.1) or intensity (-6 dB, probability = 0.1). TMS was applied over the lip representation in the left motor cortex in both Experiments 3 and 4. Left: Grand-average MMN responses to duration increments at FCz electrode. Right: Grand-average MMN responses to intensity decrements at FCz electrode. (A) Effect of TMS on responses elicited by changes in a sequence of "da" sounds in Experiment 3. (B) Effect of TMS on MMN responses elicited by changes in a sequence of piano tones (middle C) in Experiment 4. The gray areas indicate the time periods during which the baseline (black) and post-TMS (red) responses differed significantly from each other (sequential *t*-tests). (C) The mean amplitudes (\pm standard error) of MMN responses (see Table 1). Paired *t*-tests were used in statistical comparisons. **P* < 0.05, ***P* < 0.01 (2-tailed).

P < 0.05). The TMS-induced disruption of the lip representation significantly suppressed MMN responses to intensity changes in speech (*P* < 0.01), but not in tones.

Discussion

We investigated whether the human articulatory motor cortex contributes to the neural mechanisms underlying auditory discrimination. Our main finding shows that the articulatory motor cortex affects early auditory discrimination of speech sounds (starting within 200 ms after the onset of sound): TMS-induced disruption of the motor lip representation suppressed MMN responses elicited by occasional phonetic and intensity changes in unattended sequences of speech sounds. We also found evidence that the auditory system interacts

specifically with the articulatory motor cortex during speech discrimination: TMS-induced disruption of the motor hand representation did not suppress the MMN responses to changes in speech sounds. Furthermore, we found evidence of speech specificity: TMS-induced disruptions in the articulatory motor cortex had no effect on MMN responses to changes in sequences of piano tones.

MMN as an Index of Auditory Discrimination

Features of recent sounds are represented in auditory sensory memory. These auditory representations (or “memory traces”) form a prediction of future sounds. A sound that violates this prediction elicits an MMN automatically (Garrido et al. 2009; Näätänen et al. 2011). MMN can be considered to be an index of auditory discrimination, therefore. The more accurately the listener discriminates 2 nonspeech or speech sounds, the larger the MMN responses (Sams, Hämäläinen et al. 1985; Sams, Paavilainen 1985; Winkler et al. 1999; Diaz et al. 2008; Kujala and Näätänen 2010).

Previous TMS studies on the influences of motor cortex on the discrimination of speech sounds have measured reaction times and accuracy in behavioral tasks. Studies using ambiguous speech sounds have found TMS-induced effects (Meister et al. 2007; D’Ausilio et al. 2009; Möttönen and Watkins 2009), whereas studies using unambiguous speech sounds have not found such effects (Sato et al. 2009; D’Ausilio et al. forthcoming). This has been seen as a support for the view that the motor cortex contributes to speech processing in compromised listening conditions only. Our present findings are not in accord with this view. We found—using amplitude of MMN as an index of discriminability—that TMS-induced disruption in the articulatory motor cortex modulated discrimination of non-degraded natural speech sounds. The TMS-induced effects on neural processes are subtle and, therefore, sensitive measures are needed to detect their consequences. Detection of changes in discrimination of non-degraded speech sounds using simple behavioral tasks may not be possible, because of ceiling effects. Passive listening to nondegraded speech signals modulates the activity of the articulatory motor cortex (Watkins et al. 2003; Yuen et al. 2010; Murakami et al. 2011). Whether this activity is epiphenomenal or whether the articulatory motor cortex contributes to perception of clear speech remained unanswered. The present findings suggest that the articulatory motor cortex contributes to discrimination of natural non-degraded speech sounds.

In addition to phonetic changes (Experiment 1), TMS-induced disruption in the articulatory motor cortex affected MMN responses to changes in intensity, but not to changes in duration, in a sequence of “da” sounds (Experiment 3). Decreasing the intensity of consonant–vowel syllables increases their phonetic ambiguity, whereas lengthening their duration does not affect their phonetic ambiguity in English. These results, therefore, tentatively suggest that the articulatory motor cortex contributes to processing of speech sounds at a phonetic level. Importantly, the MMN responses elicited by intensity and duration changes in sequences of complex tones were unaffected by TMS-induced disruption in the articulatory motor cortex. This suggests that during the processing of non-speech sounds, the auditory system does not interact with the articulatory motor cortex. Further studies

are needed to determine whether auditory processing of non-speech signals, which can be produced with the human vocal tract or are more speech-like, is modulated by the articulatory motor cortex.

Feature-Specificity and Task-Dependence of Auditory-Motor Interactions during Speech Perception

Previous behavioral TMS studies have shown that the motor representations of articulators influence performance in speech tasks in an articulatory feature-specific manner (D’Ausilio et al. 2009; Möttönen and Watkins 2009). For example, in our previous study, TMS-induced disruption of the lip representation impaired discrimination of lip- from tongue-articulated speech sounds (“ba” vs. “da” and “pa” vs. “ta”), but had no effect on discrimination of speech sounds where the critical feature does not involve articulation at the lips (“ga” vs. “da” and “ka” vs. “ga”). In the present study, TMS-induced disruption of the motor lip representation suppressed MMN responses to both lip-articulated “ba” and tongue-articulated “ga” sounds that were presented among tongue-articulated “da” sounds. In the light of our earlier findings, the non-specificity of the TMS-induced effects in present study is surprising. It is possible that this non-specificity is due to a failure to disrupt the motor lip representation without disrupting the adjacent tongue representation. We consider, however, this to be unlikely, because the procedures used to localize and disrupt the motor lip representation in this study and our previous study were identical. We propose, therefore, that the behavioral tasks can fine-tune motor contributions to speech processing and that in the absence of behavioral tasks the feature-specificity of motor contributions is reduced. Behavioral tasks force the perceivers to selectively attend to task-relevant critical features of the speech sounds and make decisions. Tasks and selective attention have been shown to strongly modulate neural processing of speech sounds in the human auditory cortex and other temporal regions (Ahveninen et al. 2006; Jääskeläinen et al. 2007; Sabri et al. 2008). Further studies are needed to determine how tasks and selective attention affect speech processing in the articulatory motor system and its interaction with the auditory system.

Why did the TMS-induced disruption of the lip representation cause a nonspecific impairment in auditory discrimination of speech sounds in the present study? We propose that this is due to interference in automatic inverse modeling of the positions and movements of articulators that modulate the shape of the vocal tract. Although the lips close the vocal tract only when “ba” is produced but not when “da” and “ga” are produced, a complete inverse model includes the states of all articulators (including those that do not move). For example, in order to be able to repeat a word articulated by someone else, it is crucial to generate a complete motor model of the articulatory sequence. Our findings support the view that generation of these motor models is automatic (i.e. it does not require that speech signal is in the focus of attention) and that they contribute to discrimination of speech sounds. During speech production, the articulatory motor system is considered to control the movements of the articulators and generate predictions (or forward models) of the sensory consequences of these movements, which are sent to sensory systems and compared with the actual sensory input (Wolpert and Flanagan 2001; Guenther 2006; Rauschecker

and Scott 2009; Hickok et al. 2011). Testing these predictions is important for monitoring and controlling of one's own speech. It has been proposed that such predictions are also generated during speech perception (Sams et al. 2005; van Wassenhove et al. 2005; Skipper et al. 2007; Hickok et al. 2011). Tightly linked feedback (i.e. generation of motor models of a speaker's articulatory movements) and forward (i.e. predicting their sensory consequences) mechanisms could help in constructing accurate motor models of a speaker's articulatory movements and aid in categorizing his/her speech sounds. We propose that, in the present study, TMS-induced disruption of the articulatory motor system interfered with the generation of accurate motor models resulting in weaker or noisier predictions being sent to the auditory system.

Focalilty of TMS-induced Disruptions

Repetitive low-frequency TMS induces a temporary disruption, that is, "virtual lesion", in the cortex directly under the coil (Chen et al. 1997; Möttönen and Watkins 2009). It is, however, difficult to estimate reliably the extent of the disrupted region (Siebner et al. 2009; Ziemann 2010; Möttönen and Watkins forthcoming). An obvious concern related to the present study is that TMS over the motor lip representation in the left hemisphere disrupted not only the target region, but also other regions, some of which generate MMN responses. The strongest generators of the MMN responses to both speech and non-speech sounds are in primary and secondary auditory cortex in the left and right hemispheres (Hari et al. 1984; Sams, Hämäläinen et al. 1985; Sams, Paavilainen 1985; Vihla et al. 2000). It has been suggested that frontal regions, especially in the right hemisphere, also contribute to MMN responses (Giard et al. 1990; Rinne et al. 2000). In line with this, patients with unilateral lesions in the left and right dorsolateral prefrontal cortex show suppressed MMN responses to frequency changes in tone sequences (Alho et al. 1994). It is, however, unlikely that TMS disrupted the frontal and temporal generators of MMN responses and, in turn, caused the suppressions found in the current study. First, the lack of suppression of MMN responses after TMS over the motor hand representation (Experiment 2) suggests that the suppression of MMN responses after TMS over the motor lip representation (Experiment 1) was not caused by a widespread TMS-induced disruption in the frontal lobe. Secondly, if TMS over the lip representation had disrupted the frontal region or temporal region generating MMN responses, it should have suppressed MMN responses to changes in all sounds, including non-speech sounds. In contrast, TMS over the motor lip representation had no effect on the MMN responses to duration changes in speech sounds (Experiment 3) and MMN responses to both intensity and duration changes in complex tones (Experiment 4). We consider the most feasible interpretation of the present findings to be that TMS over the motor lip representation disrupted the cortex directly under the coil causing interference in its interactions with the regions that are involved in discriminating speech sounds, that is, primarily with the auditory cortex.

Is Speech Perception Auditory, Motor or Auditory-Motor?

The contributions of auditory and motor processes to speech perception have been under scientific debate for decades

(Diehl et al. 2004). Many speech scientists consider speech perception to be a purely auditory process. This view is supported by the evidence that animals that are unable to produce speech sounds are able to categorize and discriminate them (Kuhl and Miller 1975; Kuhl and Padden 1983). Changes in speech sounds can even elicit MMN responses in the rodent auditory cortex (Ahmed et al. 2011). Moreover, numerous neurophysiological studies in humans have found enhanced activity in the auditory regions in the temporal lobes during speech sound processing (Näätänen et al. 1997; Binder et al. 2000; Vihla et al. 2000; Dehaene-Lambertz et al. 2005; Parviainen et al. 2005). On the other hand, growing experimental evidence shows that the human motor system underlying speech production is activated during speech perception (Fadiga et al. 2002; Watkins et al. 2003; Wilson et al. 2004; Pulvermüller et al. 2006; Yuen et al. 2010; Murakami et al. 2011) and contributes to performance in speech tasks (Meister et al. 2007; D'Ausilio et al. 2009; Möttönen and Watkins 2009). This has lent support to the view that the motor processes play a role in both speech production and perception (Lieberman et al. 1967; Liberman and Mattingly 1985; Davis and Johnsrude 2007; Wilson 2009; Pulvermüller and Fadiga 2010; Möttönen and Watkins forthcoming). The present study provides new insight into this controversy by demonstrating that in the human brain, the articulatory motor processes interact causally with the auditory speech processing. This auditory-motor interaction facilitates processing of speech sounds and improves their discriminability even in the absence of behavioural tasks.

Funding

RM was supported by Osk Huttunen Foundation, Academy of Finland, Alfred Kordelin Foundation, Wellcome Trust and Medical Research Council. RD was supported by the Wellcome Trust. TMS equipment was funded by a grant from the John Fell Oxford University Press Fund. Funding to pay the Open Access publication charges for this article was provided by the Wellcome Trust.

Notes

We thank Prof. Dorothy Bishop for the use of EEG equipment and useful discussions and Dr Mervyn Hardiman for technical help. We thank Dr Ingrid Johnsrude for her comments on the manuscript. We also thank Rowan Boyles, Jennifer Chesters and Julia Erb for help with data collection. *Conflict of Interest:* None declared

References

- Ahmed M, Mällo T, Leppänen PHT, Hämäläinen J, Äyräväinen L, Ruusuvirta T, Astikainen P. 2011. Mismatch brain response to speech sound changes in rats. *Front Psychol.* 2:283.
- Ahveninen J, Jääskeläinen IP, Raij T, Bonmassar G, Devore S, Hämäläinen M, Levänen S, Lin FH, Sams M, Shinn-Cunningham BG et al. 2006. Task-modulated 'what' and 'where' pathways in human auditory cortex. *Proc Natl Acad Sci USA.* 103:14608–14613.
- Alho K, Woods DL, Algazi A, Knight RT, Näätänen R. 1994. Lesions of frontal cortex diminish the auditory mismatch negativity. *Electroencephalogr Clin Neurophysiol.* 91:353–362.
- Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN, Possing ET. 2000. Human temporal lobe activation by speech and nonspeech sounds. *Cereb Cortex.* 10:512–528.

- Chen R, Classen J, Gerloff C, Celnik P, Wassermann EM, Hallett M, Cohen IG. 1997. Depression of motor cortex excitability by low-frequency transcranial magnetic stimulation. *Neurology*. 48:1398–1403.
- Cheour M, Ceponiene R, Lehtokoski A, Luuk A, Allik J, Alho K, Näätänen R. 1998. Development of language-specific phoneme representations in the infant brain. *Nat Neurosci*. 1:351–353.
- Cheour M, Martynova O, Näätänen R, Erkkola R, Sillanpää M, Kero P, Raz A, Kaipio ML, Hiltunen J, Aaltonen O *et al.* 2002. Speech sounds learned by sleeping newborns. *Nature*. 415:599–600.
- D'Ausilio A, Bufalari I, Salmas P, Fadiga L. forthcoming. The role of the motor system in discriminating normal and degraded speech sounds. *Cortex*.
- 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. *Hearing 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.
- Diaz B, Baus C, Escera C, Costa A, Sebastian-Galles N. 2008. Brain potentials to native phoneme discrimination reveal the origin of individual differences in learning the sounds of a second language. *Proc Natl Acad Sci USA*. 105:16083–16088.
- Diehl RL, Lotto AJ, Holt LL. 2004. Speech perception. *Annu Rev Psychol*. 55:149–179.
- 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.
- Fischer C, Morlet D, Bouchet P, Luaute J, Jourdan C, Salord F. 1999. Mismatch negativity and late auditory evoked potentials in comatose patients. *Clin Neurophysiol*. 110:1601–1610.
- Galantucci B, Fowler CA, Turvey MT. 2006. The motor theory of speech perception reviewed. *Psychon Bull Rev*. 13:361–377.
- Garrido MI, Kilner JM, Stephan KE, Friston KJ. 2009. The mismatch negativity, a review of underlying mechanisms. *Clin Neurophysiol*. 120:453–463.
- Giard MH, Perrin F, Pernier J, Bouchet P. 1990. Brain generators implicated in the processing of auditory stimulus deviance, a topographic event-related potential study. *Psychophysiology*. 27:627–640.
- Guenther FH. 2006. Cortical interactions underlying the production of speech sounds. *J Commun Disord*. 39:350–365.
- Guthrie D, Buchwald JS. 1991. Significance testing of difference potentials. *Psychophysiology*. 28:240–244.
- Hari R, Hämäläinen M, Ilmoniemi R, Kaukoranta E, Reinikainen K, Salminen J, Alho K, Näätänen R, Sams M. 1984. Responses of the primary auditory cortex to pitch changes in a sequence of tone pips, neuromagnetic recordings in man. *Neurosci Lett*. 50:127–132.
- Hickok G, Houde J, Rong F. 2011. Sensorimotor integration in speech processing, computational basis and neural organization. *Neuron*. 69:407–422.
- Jääskeläinen IP, Ahveninen J, Belliveau JW, Raij T, Sams M. 2007. Short-term plasticity in auditory cognition. *Trends Neurosci*. 30:653–661.
- Jacobsen T, Schröger E. 2001. Is there pre-attentive memory-based comparison of pitch?. *Psychophysiology*. 38:723–727.
- Jacobsen T, Schröger E. 2003. Measuring duration mismatch negativity. *Clin Neurophysiol*. 114:1133–1143.
- Kuhl PK, Miller JD. 1975. Speech perception by the chinchilla, voiced-voiceless distinction in alveolar plosive consonants. *Science*. 190:69–72.
- Kuhl PK, Padden DM. 1983. Enhanced discriminability at the phonetic boundaries for the place feature in macaques. *J Acoust Soc Am*. 73:1003–1010.
- Kujala T, Näätänen R. 2010. The adaptive brain: a neurophysiological perspective. *Prog Neurobiol*. 91:55–67.
- Kujala T, Tervaniemi M, Schröger E. 2007. The mismatch negativity in cognitive and clinical neuroscience: theoretical and methodological considerations. *Biol Psychol*. 74:1–19.
- Lieberman AM, Cooper FS, Shankweiler DP, Studdert-Kennedy M. 1967. Perception of the speech code. *Psychol Rev*. 74:431–461.
- Lieberman AM, Mattingly IG. 1985. The motor theory of speech perception revised. *Cognition*. 21:1–36.
- Lotto AJ, Hickok GS, Holt LL. 2009. Reflections on mirror neurons and speech perception. *Trends Cogn Sci*. 13:110–114.
- 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, Watkins KE. 2009. Motor representations of articulators contribute to categorical perception of speech sounds. *J Neurosci*. 29:9819–9825.
- Möttönen R, Watkins KE. forthcoming. Using TMS to study the role of the articulatory motor system in speech perception. *Aphasiology*.
- 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.
- Näätänen R, Kujala T, Winkler I. 2011. Auditory processing that leads to conscious perception, a unique window to central auditory processing opened by the mismatch negativity and related responses. *Psychophysiology*. 48:4–22.
- Näätänen R, Lehtokoski A, Lennes M, Cheour M, Huotilainen M, Iivonen A, Vainio M, Alku P, Ilmoniemi RJ, Luuk A *et al.* 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*. 385:432–434.
- Näätänen R, Tervaniemi M, Sussman E, Paavilainen P, Winkler I. 2001. 'Primitive intelligence' in the auditory cortex. *Trends Neurosci*. 24:283–288.
- Parviainen T, Helenius P, Salmelin R. 2005. Cortical differentiation of speech and nonspeech sounds at 100 ms, implications for dyslexia. *Cereb Cortex*. 15:1054–1063.
- Pulvermüller F, Fadiga L. 2010. Active perception, sensorimotor circuits as a cortical basis for language. *Nat Rev Neurosci*. 11:351–360.
- 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 USA*. 103:7865–7870.
- Rauschecker JP, Scott SK. 2009. Maps and streams in the auditory cortex, nonhuman primates illuminate human speech processing. *Nat Neurosci*. 12:718–724.
- Rinne T, Alho K, Ilmoniemi RJ, Virtanen J, Näätänen R. 2000. Separate time behaviors of the temporal and frontal mismatch negativity sources. *Neuroimage*. 12:14–19.
- Sabri M, Binder JR, Desai R, Medler DA, Leitl MD, Liebenthal E. 2008. Attentional and linguistic interactions in speech perception. *Neuroimage*. 39:1444–1456.
- Sams M, Hämäläinen M, Antervo A, Kaukoranta E, Reinikainen K, Hari R. 1985. Cerebral neuromagnetic responses evoked by short auditory stimuli. *Electroencephalogr Clin Neurophysiol*. 61:254–266.
- Sams M, Möttönen R, Sihvonen T. 2005. Seeing and hearing others and oneself talk. *Brain Res Cogn Brain Res*. 23:429–435.
- Sams M, Paavilainen P, Alho K, Näätänen R. 1985. Auditory frequency discrimination and event-related potentials. *Electroencephalogr Clin Neurophysiol*. 62:437–448.
- Sato M, Tremblay P, Gracco VL. 2009. A mediating role of the premotor cortex in phoneme segmentation. *Brain Lang*. 111:1–7.
- Scott SK, McGettigan C, Eisner F. 2009. A little more conversation, a little less action-candidate roles for the motor cortex in speech perception. *Nat Rev Neurosci*. 10:295–302.
- Siebner HR, Hartwigsen G, Kassuba T, Rothwell JC. 2009. How does transcranial magnetic stimulation modify neuronal activity in the brain? Implications for studies of cognition. *Cortex*. 45:1035–1042.
- Skipper JI, van Wassenhove V, Nusbaum HC, Small SL. 2007. Hearing lips and seeing voices, how cortical areas supporting speech

- production mediate audiovisual speech perception. *Cereb Cortex*. 17:2387–2399.
- van Wassenhove V, Grant KW, Poeppel D. 2005. Visual speech speeds up the neural processing of auditory speech. *Proc Natl Acad Sci USA*. 102:1181–1186.
- Vihla M, Lounasmaa OV, Salmelin R. 2000. Cortical processing of change detection, dissociation between natural vowels and two-frequency complex tones. *Proc Natl Acad Sci USA*. 97:10590–10594.
- 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. 2009. Speech perception when the motor system is compromised. *Trends Cogn Sci*. 13:329–330.
- Wilson SM, Saygin AP, Sereno MI, Iacoboni M. 2004. Listening to speech activates motor areas involved in speech production. *Nat Neurosci*. 7:701–702.
- Winkler I, Kujala T, Tiitinen H, Sivonen P, Alku P, Lehtokoski A, Czigler I, Csépe V, Ilmoniemi RJ, Näätänen R. 1999. Brain responses reveal the learning of foreign language phonemes. *Psychophysiology*. 36:638–642.
- Wolpert DM, Flanagan JR. 2001. Motor prediction. *Curr Biol*. 11:729–732.
- Yuen I, Davis MH, Brysbaert M, Rastle K. 2010. Activation of articulatory information in speech perception. *Proc Natl Acad Sci USA*. 107:592–597.
- Ziemann U. 2010. TMS in cognitive neuroscience, virtual lesion and beyond. *Cortex*. 46:124–127.