

Attention Fine-Tunes Auditory–Motor Processing of Speech Sounds

Riikka Möttönen,^{1,2} Gido M. van de Ven,¹ and Kate E. Watkins¹

¹Department of Experimental Psychology and ²Oxford Centre for Human Brain Activity, Warneford Hospital, University of Oxford, Oxford, OX1 3UD, United Kingdom

The earliest stages of cortical processing of speech sounds take place in the auditory cortex. Transcranial magnetic stimulation (TMS) studies have provided evidence that the human articulatory motor cortex contributes also to speech processing. For example, stimulation of the motor lip representation influences specifically discrimination of lip-articulated speech sounds. However, the timing of the neural mechanisms underlying these articulator-specific motor contributions to speech processing is unknown. Furthermore, it is unclear whether they depend on attention. Here, we used magnetoencephalography and TMS to investigate the effect of attention on specificity and timing of interactions between the auditory and motor cortex during processing of speech sounds. We found that TMS-induced disruption of the motor lip representation modulated specifically the early auditory-cortex responses to lip-articulated speech sounds when they were attended. These articulator-specific modulations were left-lateralized and remarkably early, occurring 60–100 ms after sound onset. When speech sounds were ignored, the effect of this motor disruption on auditory-cortex responses was nonspecific and bilateral, and it started later, 170 ms after sound onset. The findings indicate that articulatory motor cortex can contribute to auditory processing of speech sounds even in the absence of behavioral tasks and when the sounds are not in the focus of attention. Importantly, the findings also show that attention can selectively facilitate the interaction of the auditory cortex with specific articulator representations during speech processing.

Key words: attention; auditory cortex; MEG; sensorimotor; speech; TMS

Introduction

Fine motor control of the articulatory movements is required for the production of speech sounds. For example, voiced stop consonants are produced by temporarily closing the vocal tract with the lips (“b”), the tip of the tongue (“d”), or the root of the tongue (“g”). It is debated whether speech perception relies on internal transformation of speech signals to articulatory movements (Lotto et al., 2009; Scott et al., 2009; Hickok, 2010; Pulvermüller and Fadiga, 2010; Möttönen and Watkins, 2012), as suggested by the motor theory of speech perception (Lieberman and Mattingly, 1985). This controversial view is supported by evidence showing that, in addition to the auditory cortex, the areas in the left motor cortex that control movements of the lips and tongue can be activated during listening to speech sounds in an articulator-specific manner (Fadiga et al. 2002; Pulvermüller et al., 2006). Moreover, transcranial magnetic stimulation (TMS) studies

demonstrate that stimulation of these motor areas affects performance in demanding speech perception tasks (Meister et al., 2007; D’Ausilio et al., 2009; Möttönen and Watkins, 2009; Sato et al., 2009). For example, TMS-induced disruption in the motor lip area impairs performance in tasks that involve discrimination of lip- and tongue-articulated sounds (e.g., “ba” and “da”) but has no effect on tasks that involve only tongue-articulated sounds (e.g., “ga” and “da”; Möttönen and Watkins, 2009).

There is evidence that the motor activations during speech perception are automatic (Chevillet et al., 2013; Möttönen et al., 2013). Recently, by combining TMS with electroencephalography we showed that the articulatory motor cortex contributes to automatic discrimination of speech sounds (Möttönen et al., 2013). We recorded mismatch negativity (MMN) responses to changes in speech sounds while participants watched a silent film, i.e., attention was directed away from the speech sounds. TMS-induced disruption of the motor lip representation suppressed MMN responses elicited by both lip-articulated “ba” and tongue-articulated “ga” sounds that were presented among tongue-articulated “da” sounds, showing that the auditory and motor cortices interact during processing of unattended speech sounds. However, intriguingly, the effect of the motor disruption on auditory speech processing was nonspecific. This finding is in conflict with the previously described articulator-specific effects. We hypothesized that the articulator-specific effects are dependent on attention and, therefore, present only during tasks that force the perceivers to focus attention on speech sounds.

Received May 24, 2013; revised Jan. 10, 2014; accepted Feb. 11, 2014.

Author contributions: R.M. and K.E.W. designed research; R.M. and G.M.v.d.V. performed research; R.M. and G.M.v.d.V. analyzed data; R.M., G.M.v.d.V., and K.E.W. wrote the paper.

This study was funded by a Medical Research Council Career Development Fellowship (R.M.).

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

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

DOI:10.1523/JNEUROSCI.2214-13.2014

Copyright © 2014 Möttönen et al.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

In the present study, we investigated how the articulatory motor cortex interacts with the auditory cortex in the different stages of speech processing and whether focusing attention on the phonetic features of speech sounds modulates the timing and articulator specificity of these auditory–motor interactions. We used magnetoencephalography (MEG) to track processing of “ba,” “da,” and “ga” sounds in the auditory cortex during TMS-induced disruption of the motor lip representation (“post-TMS” session) and in the absence of motor disruption (“no-TMS” session). The participants either performed a one-back task that forced them to attend to the phonetic features of the sounds (“attend” condition) or had no task (“ignore” condition).

Materials and Methods

Participants. Fifteen (19–31 years old, six females) right-handed native English speakers participated in the study. MEG data of three female participants were excluded because of large artifacts after TMS, which were probably caused by TMS-induced magnetization of eye makeup. MEG data of 12 participants (20–31 years old, three females) were included in the data analyses.

Stimulus sequences. The same speech stimuli as in our previous study were used (Möttönen et al., 2013). The syllables “ba,” “da,” and “ga” were produced by a female native speaker of British English. The syllables had equal duration (100 ms) and intensity. The syllables differed acoustically from each other during only the first 26 ms (i.e., during the formant transitions). The steady-state parts of the stimuli corresponding to the vowel sound (i.e., the last 74 ms) were identical in all three stimuli.

During MEG measurements, two sound sequences were presented through insert earphones at a comfortable intensity. The sequences consisted of alternating “ba” and “da” sounds (sequence 1) and “ga” and “da” sounds (sequence 2). Each sequence included 12 targets, i.e., repeated syllables. The duration of each sequence was 3 min (182 syllables with stimulus onset asynchrony of 1001 ms).

Experiment. The experiment included two MEG sessions: (1) one baseline session in the absence of any TMS-induced disruptions (no-TMS session); and (2) one immediately after the application of a 15 min train of low-frequency repetitive TMS (rTMS) over the left motor lip area (post-TMS session). The order of no-TMS and post-TMS sessions was counterbalanced. Five of 12 participants started the experiment with the no-TMS session, followed by application of rTMS and the post-TMS MEG session. Seven participants started with rTMS and the post-TMS MEG session. In these participants, the no-TMS session was performed after a 30 min break (after the end of the post-TMS MEG session) during which the motor cortex recovered from the stimulation.

Each MEG session consisted of two conditions: attend and ignore. In the attend condition, the participants were asked to attend to the sounds and to give a response by lifting the left thumb when a syllable was repeated. This response was detected using an optical response pad. The participants were asked to do this as quickly and accurately as possible. In the ignore condition, the participants were told to ignore the sounds and to focus on watching a silent film (a nature documentary). The silent movie was also presented during the attend condition. The order of the ignore and attend conditions and of the sound sequences within the condition was counterbalanced across participants.

Behavioral data. To test whether the participants were able to detect the targets (i.e., the repeated syllables) reliably, we calculated the d' value for each sequence in each session. Because of technical problems, the behavioral data of the first participant were not stored, and, therefore, data of 11 participants were included in the analyses. The one-sample t tests revealed that the participants' d' values differed from 0 in both the no-TMS (sequence 1, 2.72 ± 0.37 ; sequence 2, 3.37 ± 0.24) and post-TMS (sequence 1, 2.53 ± 0.28 ; sequence 2, 3.37 ± 0.33) sessions (all p values < 0.001). This confirms that the participants focused their attention on the distinctive features of the syllables and were able to discriminate the sounds from each other in the attend conditions. Two-way repeated-measures ANOVA showed no main effect of TMS or sequence. The interaction was also nonsignificant. This shows that TMS-induced motor disruptions had no effect on task performance. This was expected

because the syllables were unambiguous natural speech sounds. In our previous study, we showed that the TMS-induced disruption of the motor lip representation impairs discrimination of synthetic “ba” and “da” sounds that are close to the category boundary (Möttönen and Watkins, 2009). The speech sounds used in the current study were natural speech sounds and, therefore, differed acoustically more from each other than the synthetic speech sounds that are close to the category boundary. The effects of subtle TMS-induced disruptions cannot be detected using behavioral tasks if the performance is close to ceiling, as in the current study.

TMS. All TMS pulses were monophasic, generated by two Magstim 200 stimulators and delivered through a 70 mm figure-eight coil connected through a BiStim module (Magstim). The same procedure to localize the motor lip area and deliver the stimulation was used as in our previous studies (Möttönen and Watkins, 2009; Möttönen et al., 2013). We determined the active motor threshold of each participant by measuring motor-evoked potentials from the right side of the lip (orbicularis oris) muscle that was slightly contracted. The mean active motor threshold for the lip area was $57 \pm 6\%$ of the maximum intensity of the stimulator. These active motor thresholds were used as intensity for the low-frequency (0.6 Hz) rTMS, delivered for 15 min over the lip area. This rTMS train suppresses the excitability of the lip motor cortex for at least 15 min after the end of the train (Möttönen and Watkins, 2009). The coil was held in place manually. It was replaced after 7.5 min to avoid overheating. Throughout application of rTMS, the lip muscles were relaxed. No motor-evoked potentials were elicited during rTMS.

MEG recordings. Cortical magnetic signals were recorded with a 306-channel whole-head neuromagnetometer (Elekta Neuromag) in the Oxford Centre for Human Brain Activity. During data acquisition, the recorded signals were bandpass filtered at 0.03–200 Hz and digitized at 600 Hz. Horizontal and vertical eye movements were monitored with electro-oculography (EOG) recordings. Before the experiment, the positions of four marker coils, placed on the scalp, were determined in relation to three anatomical landmark points (the nasion and both preauricular points) with a 3D digitizer. This procedure allowed alignment of the MEG and MRI coordinate systems. Anatomical T1-weighted MRIs were obtained with a 3 T scanner in the Oxford Centre for Functional MRI of the Brain. The MRIs of one participant were used to visualize the locations of sources of auditory-evoked fields. Epochs of 420 ms including a prestimulus baseline of 120 ms were averaged for each stimulus in each condition and session. The MEG epochs for the targets were excluded from averaging. Also, epochs coinciding with blinks and excessive eye movements were excluded. The artifact-rejection limits were set to 5000 fT/cm for MEG channels and to 200 μ V for EOG channels. In each condition and session, ~ 70 epochs for each syllable were averaged. These averaged MEG signals were low-pass filtered at 40 Hz.

Source analysis. To obtain estimates of the time courses of neural activity in the left and right auditory cortex, we used equivalent current dipole (ECD) modeling (Hämäläinen et al., 1993). The head was modeled as a spherically symmetric conductor. For each participant, we modeled one ECD in each hemisphere during a strong dipolar magnetic field at a latency of ~ 120 ms after sound onset (i.e., peak of N100m). To increase the signal-to-noise ratio, the 3D locations and orientations of the ECDs were modeled for each participant using the MEG signals that were averaged across all syllables in the no-TMS ignore condition. Approximately 280 epochs were averaged in each participant. Both left- and right-hemisphere ECDs could be estimated reliably, with an average goodness-of-fit value of 93%.

Then, the analysis was extended to cover the entire time period (from -120 to 300 ms) and all sensors. The strengths of the ECDs were allowed to change as a function of time to best explain the evoked fields to each syllable in each condition and session. The orientations and locations of the ECDs were kept fixed. These source waveforms provide estimates of auditory-cortex activity as a function of time. It should be noted that, because only one dipole with fixed orientation was used to model the activity in each auditory cortex, the model has its limitations. For example, it cannot indicate which subregions of the auditory cortex were activated at each time point.

Next, windows of interest were defined on the basis of grand-average source waveforms (across all participants, syllables, conditions, and ses-

sions). The following windows, centered at the peaks of the auditory-evoked responses, were defined: 60–100 ms (P50m), 110–150 ms (N100m), 170–210 ms (P200m), and 220–270 ms (N250m). We then calculated mean source strengths in all of these windows for each participant, syllable, condition, and session. To test whether TMS-induced disruptions had an effect on the strengths of auditory-cortex activity in each time window, we performed five-way repeated-measures ANOVAs with the factors TMS (no-TMS vs post-TMS session), attention (attend vs ignore condition), sequence (1 vs 2), syllable (syllable 1 “ba”/“ga” vs syllable 2 “da”), and hemisphere (left vs right). Pairwise *t* tests were used in planned comparisons (two-tailed).

Results

Speech sounds elicited robust responses in the left and right auditory cortex in both the attend and ignore conditions (Fig. 1). The first response was observed 60–100 ms after sound onset (P50m), which was followed by responses at 110–150 ms (N100m), 170–210 ms (P200m), and 220–270 ms (N250m).

TMS-induced disruption of the motor lip area increased the response elicited by lip-articulated “ba” sounds at 60–100 ms when the features of speech sounds were attended (post-TMS vs no-TMS, $t_{(11)} = 6.48$, $p = 0.027$; Fig. 2*a*). This modulation was left lateralized and specific to the attend condition (TMS \times hemisphere \times attention, $F_{(1,11)} = 4.85$, $p = 0.050$; Fig. 2*b*). The modulation was also specific to “ba” sounds (sequence \times syllable \times TMS \times hemisphere \times attention, $F_{(1,11)} = 4.98$, $p = 0.047$; Fig. 2*c,d*). Responses to tongue-articulated “da” and “ga” syllables were not affected by the disruption of the motor lip representation (no significant main effect of TMS or interactions involving TMS).

The motor disruption had no effect on responses to any of the speech sounds at 110–150 ms (no significant main effect of TMS or interactions involving TMS). At 170–210 and 220–270 ms, the motor disruption modulated the responses to all the speech sounds (no significant interactions involving syllable or sequence). This modulation occurred bilaterally and during both the attend and ignore conditions at 170–210 ms (main effect of TMS, $F_{(1,11)} = 7.46$, $p = 0.019$; Fig. 3). At 220–270 ms, the TMS-induced disruption suppressed activity in the left hemisphere in the attend condition (post-TMS vs no-TMS, $t_{(11)} = 7.83$, $p = 0.017$; TMS \times hemisphere, $F_{(1,11)} = 5.90$, $p = 0.033$) but had no significant effect in the ignore condition (TMS \times hemisphere \times attention, $F_{(1,11)} = 5.12$, $p = 0.045$; Fig. 3).

Discussion

We used TMS and MEG to track dynamic interactions between the auditory and articulatory motor cortices during processing of attended and unattended speech sounds. We found evidence of two types of auditory–motor interactions: (1) early articulator-specific interactions that were dependent on attention; and (2) late nonspecific interactions that were automatic.

Early articulator-specific auditory–motor interactions

TMS-induced disruption of the lip representation increased left-hemisphere responses to attended “ba” sounds 60–100 ms after

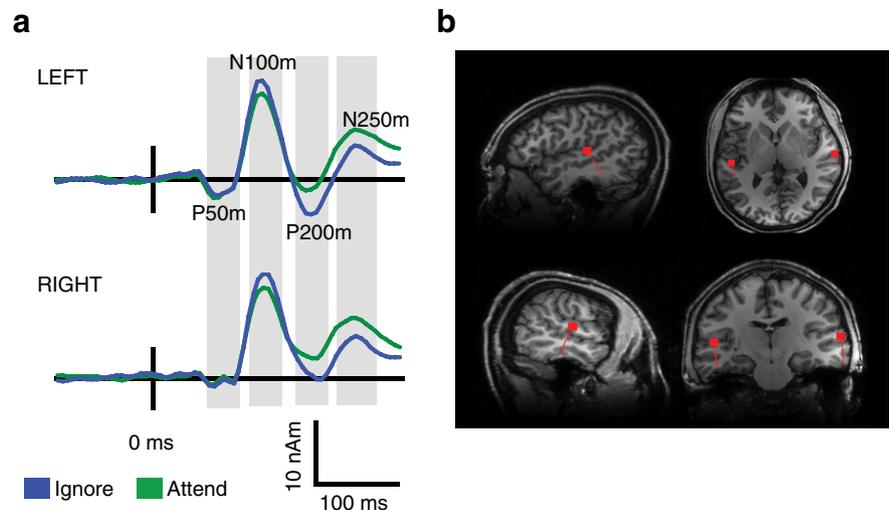


Figure 1. Time courses of activity in the left and right auditory cortices. *a*, Mean strengths of the auditory-cortex sources ($n = 12$) as a function of time during processing of all sounds (“ba,” “da,” and “ga”) in the ignore and attend conditions. Attention did not modulate activity during the first two time windows (60–100 and 110–150 ms). In later time windows, activity was modulated by attention (main effects of attention: 170–210 ms, $F_{(1,11)} = 7.49$, $p < 0.01$; 220–270 ms, $F_{(1,11)} = 16.80$, $p < 0.01$). *b*, The location of the sources, modeled as current dipoles at the peak latency of N100m, in one participant superimposed on her MRI scan. In both hemispheres, these sources are located in the superior temporal cortex in the vicinity of the Heschl’s gyrus.

sound onset. The spatiotemporal characteristics of the increased response are similar to the P50m (i.e., magnetic P1/P50) response, which is particularly prominent during auditory processing in infancy and suppressed during maturation (Sharma et al., 1997; Sussman et al., 2008), perhaps because of the development of connections between the auditory cortex and other cortical areas, such as the speech motor system. Sensory-gating studies demonstrate that the P50m response is increased in schizophrenics who experience auditory hallucinations (Smith et al., 2013) and in stutterers (Kikuchi et al., 2011), possibly reflecting abnormal corticocortical interactions. Thus, the increased P50m observed here might be a sign of reduced efficiency of auditory processing attributable to disrupted cortical interactions. We propose that the articulatory motor cortex normally inhibits the generators of the P50m response to speech sounds in the left auditory cortex, and, therefore, the amplitude of the response was increased during TMS-induced motor disruption.

It is likely that, in the earliest stage of cortical speech processing (<100 ms), acoustic–phonetic features of speech sounds are extracted from speech sounds (Tavabi et al., 2007). In the current study, the onsets of the syllables “ba,” “da,” and “ga” differed from each other acoustically, which may explain why the articulator-specific effect was observed at such an early latency. The findings suggest that focusing attention on the place-of-articulation features facilitates interactions between specific motor representations and the auditory cortex. Thus, when features of lip-articulated “ba” sounds were attended, the interaction between the motor lip representation and the auditory cortex was facilitated. Consequently, the disruption of the motor lip representation modulated the early auditory processing of attended “ba” sounds but had no effect on the early processing of other speech sounds. This finding suggests that attention can facilitate auditory–motor processing of speech sounds in an articulator-specific manner.

Late nonspecific auditory–motor interactions

TMS-induced disruption of the motor lip representation modulated responses to both lip- and tongue-articulated speech

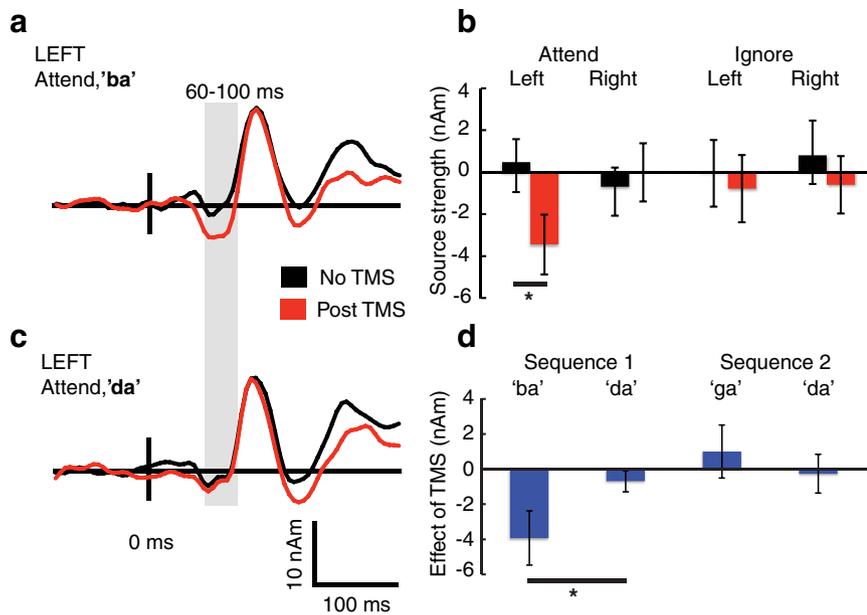


Figure 2. TMS-induced disruption of the lip representation modulated auditory-cortex responses to attended “ba” sounds in the left hemisphere at 60–100 ms. *a*, Time courses of auditory-cortex activity for attended “ba” sounds in the left hemisphere in the no-TMS (black) and post-TMS (red) sessions ($n = 12$). *b*, Mean source strengths in the left and right hemispheres at 60–100 ms for attended and ignored “ba” sounds. *c*, Time courses of left auditory-cortex activity for attended “da” sounds in the no-TMS and post-TMS sessions. *d*, Differences in the strengths of the left auditory-cortex sources between the post-TMS and no-TMS sessions for all sounds in the attend condition at 60–100 ms. TMS-induced disruption of the lip representation affected specifically early responses to “ba” sounds but not “da” sounds in sequence 1 ($t_{(11)} = -2.45, p = 0.032$). TMS had no effect on early responses to “ga” and “da” sounds in sequence 2. $*p < 0.05$ (paired t test, two-tailed). Error bars indicate SEM.

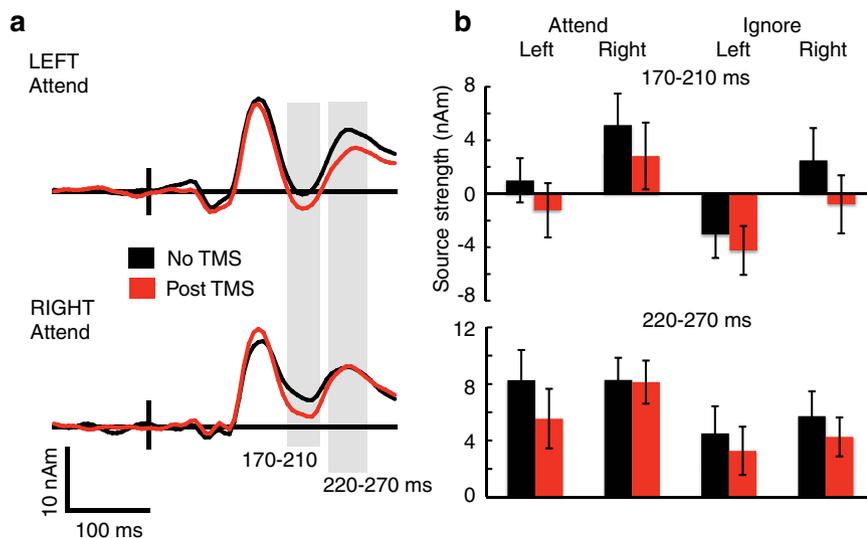


Figure 3. TMS-induced disruption of the lip representation modulated auditory-cortex responses to all speech sounds at 170–210 and 220–270 ms. *a*, Time courses of activity in the left and right auditory cortices averaged across all sounds (“ba,” “da,” and “ga”) in the attend condition in the no-TMS (black) and post-TMS (red) sessions ($n = 12$). *b*, Mean strengths of left and right auditory-cortex sources in the attend and ignore conditions at 170–210 and 220–270 ms. Error bars indicate SEM.

sounds starting 170 ms after syllable onset. The finding provides additional evidence of the automaticity of the involvement of the articulatory motor cortex in processing of speech sounds (Chevillet et al., 2013; Möttönen et al., 2013). The lack of articulator specificity of these late effects is in agreement with our previous study, which showed that disruption of the lip representation suppresses MMN responses peaking ~180 ms after the onset of unattended “ba” and “ga” sounds presented among “da” sounds

(Möttönen et al., 2013). Thus, our current and previous findings show that, from ~170 ms, the lip motor cortex modulates auditory processing of both attended and unattended speech sounds regardless of how they are produced. This lack of articulator specificity suggests that, in this phonological stage, the positions of all articulators are modeled. Interestingly, the late effects were bilateral, suggesting that the left articulatory motor cortex interacts with both left and right auditory cortices during phonological processing.

The articulatory motor cortex and speech perception

During speech perception, the articulatory motor cortex is activated (Fadiga et al., 2002; Wilson et al., 2004, 2008; Pulvermüller et al., 2006) and functionally connected to auditory regions (Wilson and Iacoboni, 2006). Furthermore, stimulation of the motor areas can affect performance in demanding speech perception tasks (D’Ausilio et al., 2009; Möttönen and Watkins, 2009). It is debatable whether the motor cortex contributes to speech perception or to postperceptual processes, such as decision-making and response selection (Hickok, 2010; Venezia et al., 2012). It can be argued that, if the articulatory motor cortex contributed to postperceptual processing, but not to speech perception, motor disruptions should have no effect on sensory processing of speech sounds in the auditory cortex. Our current and previous findings (Möttönen et al., 2013) show that this is not the case. Also, we show that, although behavioral tasks can modulate the auditory–motor interactions, these interactions occur even in the absence of behavioral tasks and when attention is directed away from the speech sounds.

Neuroanatomical models

The articulatory motor cortex is thought to contribute to speech perception by internally simulating a speaker’s articulatory movements (Stevens and Halle, 1967; Liberman and Mattingly, 1985). The current findings lend support to neuroanatomical models that propose that speech perception is based on neural circuits that reciprocally connect motor and sensory systems (Pulvermüller and Fadiga, 2010). According to other neuroanatomical models of speech processing, the motor and sensory systems are segregated but linked by the auditory dorsal stream (Hickok and Poeppel, 2007; Rauschecker and Scott, 2009). This stream transforms auditory speech signals to motor codes enabling imitation and repetition of speech signals, i.e., it generates inverse models. During speech production, this stream is involved in predicting sensory consequences of articulatory

movements, i.e., it also generates forward models. The current findings suggest that the dorsal stream generates both inverse and forward models during auditory speech processing. In other words, the auditory speech signals are transformed to motor models, which in turn affect sensory processing. The findings also suggest that attention can facilitate the generation of motor models and enhance their specificity.

Predictive coding

Previous expectations can influence speech perception (Remez et al., 1981). For example, intelligibility and perceptual clarity of speech signals increase when they match expectations (Jacoby et al., 1988; Goldinger et al., 1999; Davis et al., 2005). Expectations may influence sensory processing via fronto-temporal top-down mechanisms (Davis and Johnsrude, 2007; Sohoglu et al., 2012). The stimulus sequences in the current study were highly predictable; participants could predict reliably whether the next sound would be “ba,” “da,” or “ga.” TMS-induced motor disruptions may have interfered with the generation of these predictions and top-down influences on speech processing in the auditory cortex. This interpretation is consistent with the proposal that top-down influences are associated with motor representations (Davis and Johnsrude, 2007), but it needs to be investigated in additional studies. However, previous expectations cannot completely explain the motor contributions to speech perception, because previous TMS studies presented syllables in a random order (D’Ausilio et al., 2009; Möttönen and Watkins, 2009; Möttönen et al., 2013).

Attention and speech processing

Processing of speech sounds is considered a highly automatic process (Näätänen et al., 2001, 2007). However, focusing attention on speech sounds modulates their processing in the auditory cortex (Ahveninen et al., 2006; Sabri et al., 2008; Wild et al., 2012). We investigated whether the interactions between the auditory and articulatory motor cortices are automatic or dependent on attention. The results suggest that, although the auditory and motor cortices interact during processing of unattended speech sounds (Möttönen et al., 2013), the earliest interactions depend on attention. The task used in the current study forced the participants to focus attention on the place-of-articulation features of the speech sounds. Modulations of auditory-cortex activity during speech processing are task dependent, i.e., phonetic and nonphonetic tasks modulate the activity differently (Ahveninen et al., 2006). Therefore, it is unlikely that focusing attention on nonphonetic features of sounds would have a similar articulator-specific effect as found in the current study.

Focality of TMS-induced disruptions

Low-frequency rTMS over the motor cortex induces a temporary disruption, i.e., suppresses excitability, in the motor representation directly under the coil (Chen et al., 1997; Möttönen and Watkins, 2009). However, it is challenging to estimate the extent of the disrupted region (Siebner et al. 2009; Ziemann 2010). A concern related to the current study is that TMS over the left-hemispheric motor lip representation disrupted not only the target region but also nearby and connected regions. The fact that the early effect was highly specific to the place of articulation and to the focus of the attention suggests that it was not caused by a widespread disruption in the left hemisphere. Conversely, the nonspecific late effect could in principle be caused by a widespread disruption. However, our previous results suggest that this is not the case (Möttönen et al., 2013). According to these

results, the late effect is specific to the stimulated motor representation (lip not hand) and to the stimulus material (speech not non-speech).

Conclusions

Our findings suggest that the articulatory motor cortex contributes to two stages of speech processing in the auditory cortex. In the later stage, the auditory cortex in both hemispheres interacts with the left articulatory motor cortex during processing of both attended and unattended speech sounds. In contrast, the early auditory–motor interactions are articulator specific, left lateralized, and dependent on attention. The findings support the view that interacting auditory and motor brain regions contribute to speech perception. This auditory–motor interaction can be fine-tuned by focusing attention on the phonetic features of speech sounds. Moreover, our findings demonstrate that the novel combination of TMS and MEG provides a powerful tool to investigate the timing of sensorimotor interactions.

References

- Ahveninen J, Jääskeläinen IP, Raij T, Bonmassar G, Devore S, Hämäläinen M, Levänen S, Lin FH, Sams M, Shinn-Cunningham BG, Witzel T, Belliveau JW (2006) Task-modulated “what” and “where” pathways in human auditory cortex. *Proc Natl Acad Sci U S A* 103:14608–14613. [CrossRef Medline](#)
- Chen R, Classen J, Gerloff C, Celnik P, Wasserman EM, Hallett M, Cohen LG (1997) Depression of motor cortex excitability by low-frequency transcranial magnetic stimulation. *Neurology* 48:1398–1403. [Medline](#)
- Chevillet MA, Jiang X, Rauschecker JP, Riesenhuber M (2013) Automatic phoneme category selectivity in the dorsal auditory stream. *J Neurosci* 33:5208–5215. [CrossRef Medline](#)
- 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. [CrossRef Medline](#)
- Davis MH, Johnsrude IS (2007) Hearing speech sounds: top-down influences on the interface between audition and speech perception. *Hear Res* 229:132–147. [CrossRef Medline](#)
- Davis MH, Johnsrude IS, Hervais-Adelman A, Taylor K, McGettigan C (2005) Lexical information drives perceptual learning of distorted speech: evidence from the comprehension of noise-vocoded sentences. *J Exp Psychol Gen* 134:222–241. [CrossRef Medline](#)
- Fadiga L, Craighero L, Buccino G, Rizzolatti G (2002) Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *Eur J Neurosci* 15:399–402. [CrossRef Medline](#)
- Goldinger SD, Kleider HM, Shelley E (1999) The marriage of perception and memory: creating two-way illusions with words and voices. *Mem Cognit* 27:328–338. [CrossRef Medline](#)
- Hämäläinen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV (1993) Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev Mod Phys* 65:413–497. [CrossRef](#)
- Hickok G (2010) The role of mirror neurons in speech and language processing. *Brain Lang* 112:1–2. [CrossRef Medline](#)
- Hickok G, Poeppel D (2007) The cortical organization of speech processing. *Nat Rev Neurosci* 8:393–402. [CrossRef Medline](#)
- Jacoby LL, Allan LG, Collins JC, Larwill LK (1988) Memory influences subjective experience: noise judgments. *J Exp Psychol Learn Mem Cogn* 14:240–247. [CrossRef](#)
- Kikuchi Y, Ogata K, Umetsaki T, Yoshiura T, Kenjo M, Hirano Y, Okamoto T, Komune S, Tobimatsu S (2011) Spatiotemporal signatures of an abnormal auditory system in stuttering. *Neuroimage* 55:891–899. [CrossRef Medline](#)
- Liberman AM, Mattingly IG (1985) The motor theory of speech perception revised. *Cognition* 21:1–36. [CrossRef Medline](#)
- Lotto AJ, Hickok GS, Holt LL (2009) Reflections on mirror neurons and speech perception. *Trends Cogn Sci* 13:110–114. [CrossRef Medline](#)
- Meister IG, Wilson SM, Deblieck C, Wu AD, Iacoboni M (2007) The essential role of premotor cortex in speech perception. *Curr Biol* 17:1692–1696. [CrossRef Medline](#)
- Möttönen R, Watkins KE (2009) Motor representations of articulators con-

- tribute to categorical perception of speech sounds. *J Neurosci* 29:9819–9825. [CrossRef Medline](#)
- Möttönen R, Watkins KE (2012) Using TMS to study the role of the articulatory motor system in speech perception. *Aphasiology* 26:1103–1118. [CrossRef Medline](#)
- Möttönen R, Dutton R, Watkins KE (2013) Auditory–Motor processing of speech sounds. *Cereb Cortex* 23:1190–1197. [CrossRef Medline](#)
- Näätänen R, Lehtokoski A, Lennes M, Cheour M, Huutilainen M, Iivonen A, Vainio M, Alku P, Ilmoniemi RJ, Luuk A, Allik J, Sinkkonen J, Alho K (1997) Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385:432–434. [CrossRef Medline](#)
- Näätänen R, Tervaniemi M, Sussman E, Paavilainen P, Winkler I (2001) “Primitive intelligence” in the auditory cortex. *Trends Neurosci* 24:283–288. [CrossRef Medline](#)
- Pulvermüller F, Fadiga L (2010) Active perception: sensorimotor circuits as a cortical basis for language. *Nat Rev Neurosci* 11:351–360. [CrossRef Medline](#)
- 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. [CrossRef Medline](#)
- Rauschecker JP, Scott SK (2009) Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci* 12:718–724. [CrossRef Medline](#)
- Remez RE, Rubin PE, Pisoni DB, Carrell TD (1981) Speech perception without traditional speech cues. *Science* 212:947–949. [CrossRef Medline](#)
- Sabri M, Binder JR, Desai R, Medler DA, Leitl MD, Liebenthal E (2008) Attentional and linguistic interactions in speech perception. *Neuroimage* 39:1444–1456. [CrossRef Medline](#)
- Sato M, Tremblay P, Gracco VL (2009) A mediating role of the premotor cortex in phoneme segmentation. *Brain Lang* 111:1–7. [CrossRef Medline](#)
- 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. [CrossRef Medline](#)
- Sharma A, Kraus N, McGee TJ, Nicol TG (1997) Developmental changes in P1 and N1 central auditory responses elicited by consonant-vowel syllables. *Electroencephalogr Clin Neurophysiol* 104:540–545. [Medline](#)
- 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. [CrossRef Medline](#)
- Smith DM, Grant B, Fisher DJ, Borraji G, Labelle A, Knott VJ (2013) Auditory verbal hallucinations in schizophrenia correlate with P50 gating. *Clin Neurophysiol* 124:1329–1335. [CrossRef Medline](#)
- Sohoglu E, Peelle JE, Carlyon RP, Davis MH (2012) Predictive top-down integration of prior knowledge during speech perception. *J Neurosci* 32:8443–8453. [CrossRef Medline](#)
- Stevens KN, Halle M (1967) Remarks on analysis by synthesis and distinctive features. In: *Models for perception of speech and visual form: proceedings of a symposium* (Wathen-Dunn W, ed), pp 88–102. Cambridge, MA: Massachusetts Institute of Technology.
- Sussman E, Steinschneider M, Gumenyuk V, Grushko J, Lawson K (2008) The maturation of human evoked brain potentials to sounds presented at different stimulus rates. *Hear Res* 236:61–79. [CrossRef Medline](#)
- Tavabi K, Obleser J, Dobel C, Pantev C (2007) Auditory evoked fields differentially encode speech features: an MEG investigation of the P50m and N100m time courses during syllable processing. *Eur J Neurosci* 25:3155–3162. [CrossRef Medline](#)
- Venezia JH, Saberi K, Chubb C, Hickok G (2012) Response bias modulates the speech motor system during syllable discrimination. *Front Psychol* 3:157. [CrossRef Medline](#)
- Wild CJ, Yusuf A, Wilson DE, Peelle JE, Davis MH, Johnsrude IS (2012) Effortful listening: the processing of degraded speech depends critically on attention. *J Neurosci* 32:14010–14021. [CrossRef Medline](#)
- Wilson SM, Iacoboni M (2006) Neural responses to non-native phonemes varying producibility: evidence for the sensorimotor nature of speech perception. *Neuroimage* 33:316–325. [CrossRef Medline](#)
- Wilson SM, Saygin AP, Sereno MI, Iacoboni M (2004) Listening to speech activates motor areas involved in speech production. *Nat Neurosci* 7:701–702. [CrossRef Medline](#)
- Wilson SM, Molnar-Szakacs I, Iacoboni M (2008) Beyond superior temporal cortex: intersubject correlations in narrative speech comprehension. *Cereb Cortex* 18:230–242. [CrossRef Medline](#)
- Ziemann U (2010) TMS in cognitive neuroscience, virtual lesion and beyond. *Cortex* 46:124–127. [CrossRef Medline](#)