



ELSEVIER

Contents lists available at ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

A mediating role of the premotor cortex in phoneme segmentation

Marc Sato^{a,*}, Pascale Tremblay^{b,c}, Vincent L. Gracco^{b,c,d}

^a Gipsa-Lab, UMR CNRS 5216, Département Parole et Cognition, Grenoble Universités, 1180 avenue centrale, BP 25, 38040 Grenoble Cedex 9, France

^b School of Communication Sciences and Disorders, McGill University, 1266 Pine Avenue West, Montreal, Quebec, Canada H3G 1A8

^c Centre for Research on Language, Mind and Brain, McGill University, 3640 de la Montagne, Montreal, Quebec, Canada H3G 2A8

^d Haskins Laboratories, 300 George St., Suite 900, New Haven, CT 06511, USA

ARTICLE INFO

Article history:

Accepted 15 March 2009

Available online xxxxx

Keywords:

Speech perception
Speech production
Phonological processing
Speech segmentation
Transcranial magnetic stimulation
Premotor cortex
Mirror-neuron system
Motor theory of speech perception

ABSTRACT

Consistent with a functional role of the motor system in speech perception, disturbing the activity of the left ventral premotor cortex by means of repetitive transcranial magnetic stimulation (rTMS) has been shown to impair auditory identification of syllables that were masked with white noise. However, whether this region is crucial for speech perception under normal listening conditions remains debated. To directly test this hypothesis, we applied rTMS to the left ventral premotor cortex and participants performed auditory speech tasks involving the same set of syllables but differing in the use of phonemic segmentation processes. Compared to sham stimulation, rTMS applied over the ventral premotor cortex resulted in slower phoneme discrimination requiring phonemic segmentation. No effect was observed in phoneme identification and syllable discrimination tasks that could be performed without need for phonemic segmentation. The findings demonstrate a mediating role of the ventral premotor cortex in speech segmentation under normal listening conditions and are interpreted in relation to theories assuming a link between perception and action in the human speech processing system.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

Given the non-linear mapping between phonetic segments and their acoustic realization, distinct theoretical perspectives have been proposed to explain speech perception. A crucial distinction among these perspectives is the use of purely auditory perceptual mechanisms, or that of articulatory control processes and more generally the involvement of the motor system (for reviews, see Diehl, Lotto, & Holt, 2004; Galantucci, Fowler, & Turvey, 2006; Schwartz, Abry, Boë, & Cathiard, 2002). Many researchers have advocated models of speech perception that focus only on the auditory system and on the acoustic properties of speech (for a review, see Diehl et al., 2004). It is hypothesized that the speech signal is highly structured in that it contains invariant acoustic patterns for phonetic features, and that speech sounds are perceived using the same mechanisms of audition and perceptual learning. From this view, speech perception is assumed to be driven by auditory mechanisms, based on invariant properties of the acoustic signal, and not mediated by the motor system. In contrast, in the motor theory of speech perception (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Liberman & Mattingly, 1985; Liberman & Whalen, 2000), the objects of speech perception are the speaker's intended articulatory gestures rather than acoustic or auditory events. According to this view, speech gestures are represented in the brain as invariant motor control structures

while their manifestation in the acoustic signal or in the articulatory movements may vary contextually. Since speaker and listener share a common repertoire of motor primitives, speech perception is thought as a set of perceptuo-motor processes, used to recover the articulated gestures as the objects of perception. As such, speech perception is a sensorimotor process in that perceiving speech is perceiving speech motor (articulatory) gestures.

Consistent with this view, recent brain imaging studies demonstrate that frontal brain areas involved in the planning and execution of speech gestures (i.e., the posterior part of the left inferior frontal gyrus, namely Broca's area, and the ventral premotor cortex) are activated during passive auditory, visual and/or auditory-visual speech perception (e.g., Ojanen et al., 2005; Pekkola et al., 2006; Pulvermuller et al., 2006; Skipper, Nusbaum, & Small, 2005; Skipper, Van Wassenhove, Nusbaum, & Small, 2007; Wilson & Iacoboni, 2006; Wilson, Saygin, Sereno, & Iacoboni, 2004). Single-pulse TMS studies also show that lip or tongue motor-evoked potentials (MEP) are enhanced during both passive speech listening and viewing, when stimulating the corresponding area of the left primary motor cortex (Fadiga, Craighero, Buccino, & Rizzolatti, 2002; Roy, Craighero, Fabbri-Destro, & Fadiga, 2008; Sundara, Namasivayam, & Chen, 2001; Watkins & Paus, 2004; Watkins, Strafella, & Paus, 2003). This increased excitability of the motor system during speech perception is related to an increase in activity in Broca's area and the ventral premotor cortex (vPMC), as shown by a recent study combining the technique of positron emission tomography with TMS (Watkins & Paus, 2004). Finally, this speech motor 'resonance' mechanism (Fadiga et al., 2002) appears to be

* Corresponding author.

E-mail address: marc.sato@gipsa-lab.inpg.fr (M. Sato).

articulator specific, motor facilitation being stronger when the recorded muscle activity and the auditory speech stimulus reflect the same articulator (Fadiga et al., 2002; Roy et al., 2008). This is also suggested by two recent functional magnetic resonance imaging (fMRI) studies showing similar somatotopic patterns of motor activity in the vPMC during both producing and listening to or viewing lips- and tongue-related phonemes (Pulvermuller et al., 2006; Skipper et al., 2007). Altogether, these studies support the idea that both Broca's area and the vPMC are recruited during speech processing, and suggest that speech perception involves a specific mapping from the speaker's articulatory gestures into the listener's motor plans.

However, despite accumulating evidence that passive speech perception induces cortical activity in both Broca's area, the vPMC and the orofacial primary motor cortex, whether speech motor centers actually contribute to speech perception remains debated. While previous brain imaging and single-pulse TMS studies demonstrate the recruitment of motor areas during passive speech perception, these results are intrinsically correlational and cannot be used to address causality. Two powerful techniques for establishing causality are through the use of rTMS and electrocortical stimulation during awake neurosurgical operations to directly manipulate brain functioning. Both techniques can be used to temporarily disrupt neural activity of a given cortical region, thereby creating a transient 'virtual lesion', and to explore causal relations between the stimulated region and specific motor, sensory and/or cognitive functions (for reviews, see Boatman, 2004; Devlin & Watkins, 2007; Pascual-Leone, Walsh, & Rothwell, 2000; Walsh & Cowey, 2000). However, both electrocortical mapping studies and rTMS studies are inconclusive regarding a possible functional role of Broca's area and the left svPMC in auditory speech processing under normal listening conditions. Indeed, temporarily disrupting the activity of the left inferior frontal gyrus does not impair participants' ability to perform auditory speech discrimination tasks (Boatman, 2004; Boatman & Miglioretti, 2005) but phonological judgments that likely involve verbal working memory and articulatory rehearsal (Boatman, 2004; Romero, Walsh, & Papagno, 2006). On the other hand, a recent rTMS study showed that stimulating the superior portion of the left ventral premotor cortex (svPMC), a region found to be activated during both syllable perception and production in recent fMRI studies (Wilson et al., 2004; Pulvermuller et al., 2006), impaired auditory syllable identification (Meister, Wilson, Deblieck, Wu, & Iacoboni, 2007). The results were interpreted to suggest that the premotor cortex is an essential component in speech perception and that it may contribute to it through a top-down modulation of temporal cortex. However, it should be noted that the auditory syllable identification task was performed in the presence of masking noise which reduced performance overall and therefore impacts on the interpretation of the results. Because Meister et al. (2007) did not examine the effects of premotor stimulation on speech processing in the absence of masking noise, it is still unclear whether the motor system is functionally activated under normal speech processing conditions and, if not, whether motor system involvement is only functional in the presence of sensory challenge or is activated more generally when task demands (beyond increasing signal to noise) are increased.

To extend and refine the observations from the Meister et al. (2007) study, we examined the influence of left svPMC in speech perception under normal listening conditions and a number of auditory speech tasks. We used 1 Hz low-frequency rTMS and frameless stereotaxy to temporarily disrupt the activity of svPMC and examined participants' performance on a phoneme identification task, a syllable discrimination task and a phoneme discrimination task. All these tasks involved the same set of nonsense syllables, presented without masking noise, but differed in the

use of phonemic segmentation and verbal working memory demands. We hypothesized that the phoneme identification task, in which participants were asked to judge whether a syllable began with /p/ or /b/, could be performed solely based on an acoustic-auditory analysis of the stop consonant voicing at the onset of the syllable, without the need to recruit the motor system (Boatman, 2004; Burton, Small, & Blumstein, 2000). In the syllable identification task, because the presented syllable pairs differed (or not) only in their first initial phonemes (e.g., /put/-/but/), a similar acoustic-phonetic analysis is also likely required together with minimal verbal storage to discriminate phonetic contrasts and compare the two syllables. Finally, in the phoneme discrimination task, participants were asked to determine whether the initial phonemes of syllable pairs were the same or different. Because half of the syllable pairs differed in more complex ways from just their initial segment (e.g., /put/-/bon/), it is likely that the listener had to segment the initial phonemes from the remainder of the syllables and then compare them in order to make a same/different judgment (Burton & Small, 2006; Burton et al., 2000). We hypothesized that specific effects of rTMS over the left svPMC on accuracy and reaction times should be stronger in the phoneme discrimination task, the motor system being likely recruited to assist in phonological segmentation and working memory processes (Romero et al., 2006). According to the above-mentioned theories of speech perception, this would argue against the view that speech perception relies exclusively on the auditory system and the acoustic properties of speech but also would refine a possible mediating role of the premotor cortex in speech perception under normal listening conditions in segmenting the speech stream into constituent phonemes.

2. Methods

2.1. Participants

Ten healthy adults (seven females; mean age \pm SD, 27 \pm 5 years) participated in the study. All but one were native speakers of Canadian French. The other one was a native speaker of French. All participants were right-handed (Oldfield, 1971), had normal or corrected-to-normal vision and reported no history of speaking or hearing disorders. Participants were screened for neurological, psychiatric, and other medical conditions, and contraindications to TMS (Wassermann, 1998). Written informed consent was obtained for all participants and they were paid for their participation. The protocol was approved by the McGill University Faculty of Medicine and was carried-out in accordance with the ethical standards of the 1964 Declaration of Helsinki.

2.2. Stimuli

Utterances of /put/, /but/, /pyd/, /byd/, /bon/, /pon/ consonant-vowel-consonant (CVC) syllables were individually recorded by a male native Canadian French speaker in a sound-attenuated room. The speaker produced each syllable naturally, maintaining an even intonation, tempo and vocal intensity. Audio digitizing was done at 44.1 kHz with 16 bits. All syllables were temporally aligned according to their consonantal onsets and matched for intensity.

2.3. Procedure

The experiment was carried-out in a quiet, darkened room. Participants sat comfortably in front of a computer monitor at a distance of approximately 50 cm. The acoustic stimuli were presented at a comfortable sound level through two loudspeakers. Participants underwent two identical experimental sessions fol-

lowing rTMS and sham stimulation (see below). The interval between sessions was 1 h. Each session included four tasks: a phoneme identification task, a syllable discrimination task, a phoneme discrimination task, and a nonverbal visual matching control task. The visual matching task was designed to control for any nonspecific effects of rTMS that might affect performance. The order of the tasks was fully counterbalanced across participants but was the same across the two sessions.

In all tasks, each trial started with a fixation cue (the '+' symbol) presented in the middle of the screen for 250 ms and ended with a blank screen for 2500 ms. In the control task, the fixation cue was followed by another fixation cue presented for 500 ms on either the right, for half of the trials, or the left part of the screen (at a distance from the middle of approximately 5°). In the three auditory tasks the same set of six nonsense consonant-vowel-consonant (CVC) syllables was used (i.e., /put/, /but/, /pyd/, /byd/, /bon/, /pon/). In the phoneme identification task, the fixation cue was followed by the presentation of a nonsense syllable for 500 ms, beginning with either /p/, for half of the trials, or /b/ (e.g., /put/, /bon/). In the syllable and phoneme discrimination tasks, the fixation cue was followed by the presentation of a pair of similar or dissimilar nonsense syllables for 500 ms each, both beginning either with the same or different bilabial consonants. In the syllable discrimination task, the syllable pairs differed in their initial consonants for half of the trials (e.g., /byd/-/pyd/) and were the same for the other half (e.g., /put/-/put/). In the phoneme discrimination task, the initial consonants were different for half of the pairs, with the final VC sequences similar or different (e.g., /pon/-/bon/, /bon/-/pyd/), and were the same for the other half (e.g., /put/-/put/, /but/-/byd/).

Participants were instructed to produce a motor response as fast and accurately as possible, by pressing, with either their right index or ring finger, one of two keys corresponding to either left/right position (control task), /p/ or /b/ (phoneme identification task), similar/dissimilar syllables (syllable discrimination task), or similar/dissimilar first phonemes (phoneme discrimination task). The response key designation was fully counterbalanced across participants.

Before starting the experiment, a complete training session was performed, with the order of the tasks identical to that used in the two experimental sessions. Before each task, instructions related to the task were presented on the screen for 10 s. Because of the short delay between tasks, six practice trials were included at the beginning of each task and were removed from the analysis. Following the practice trials, all tasks included 12 randomly presented trials with the exception of the phoneme discrimination task (24 trials). The total duration of the experiment was about 6 min. The experimental design (task duration, number of trials) was selected because previous experiments with comparable rTMS parameters have shown that effects of magnetic stimulation last up to 10 min (see Fitzgerald, Fountain, & Daskalakis, 2006, for a review).

2.4. MRI Localizer

Prior to the study, a high-resolution T1-weighted structural volume was acquired for anatomical localization in all participants. Data were recorded on a 1.5 T Siemens Sonata MR scanner at the Montreal Neurological Institute (matrix 256 × 256 mm, 160 slices, 1 × 1 × 1 mm, no gap, TE = 9.2 ms, TR = 22 ms).

2.5. Transcranial magnetic stimulation

TMS was applied with a 70-mm air-cooled figure-of-eight TMS coil, driven by a high-speed magnetic stimulator (Magstim Rapid 1400, Wales, UK). Resting motor threshold (RMT) was determined in the right first dorsal interosseus muscle and defined as the minimum stimulus intensity capable of evoking motor-evoked poten-

tials (MEP) in five out of 10 trials with an amplitude of at least 50 μ V (Rossini et al., 1994). Each participant then underwent two experimental sessions, consisting of either stimulations applied over the left svPMC or sham stimulation applied over the same site with identical coil position and acoustic artefacts. During both sessions, participants wore earplugs because of rTMS-related noise whereas head motion was restricted using a foam-padded head-neck cradle. The two sessions were separated by a period of 1 h to avoid plasticity effects in the excitability of the stimulated region, and were fully counterbalanced across participants. 600 pulses were applied at a frequency of 1 Hz, with a stimulation intensity at 110% of individual RMT. These stimulation parameters have been shown to produce inhibitory modulation of motor cortical excitability, lasting up to 10 min (see Fitzgerald et al., 2006 for a review). The stimulation site was located with frameless stereotaxy at the beginning of each session on the basis of each participant's anatomical MRI scan. The MRI-to-head co-registration was done using Brainsight TMS system (Rogue Research, Montreal, Canada). For all participants, we used the average peak coordinates in the left precentral gyrus found to be activated during both listening to and producing syllables in a previous study (Wilson et al., 2004). This peak was located on the anterior lip of the central sulcus on the border of areas 4a and 6 (Montreal Neurological Institute (MNI) coordinates in mm: -50, -6 and 47). Two other studies found similar peak activation in the precentral gyrus during both syllable perception and production (Pulvermuller et al., 2006: -54, -3 and 46; Meister et al., 2007: -53, -4 and 49). For all participants, the stimulation site fell in the left svPMC (see Fig. 1). None of the participants reported any side effects following stimulation. The sham stimulation was performed by presenting participants with recorded TMS noise mimicking actual stimulation (600 pulses at a frequency of 1 Hz at a similar intensity level) through loudspeakers and with participants wearing earplugs. The coil was positioned over the left svPMC using the same localization procedure as in the rTMS session but without actually presenting any TMS stimulation. Participants were not aware that one of the two sessions involved sham stimulation.

2.6. Data analysis

Outlier reaction times (RT), i.e., responses slower than 2000 ms, were considered as errors (on average 1.8% (± 3) and 2.3% (± 3) in rTMS and sham sessions, respectively). To statistically assess the effect of various experimental conditions, two-way analyses of variance (ANOVA) for repeated measures were performed on the percentage of correct responses and on the median RTs. RTs were calculated from the onset of the second fixation cue in the control task, from the onset of the presented syllable in the phoneme identification task, and from the onset of the second presented syllable in the syllable and phoneme discrimination tasks. The independent variables were the stimulation mode (rTMS, sham) and the task (control, phoneme identification, syllable discrimination, phoneme discrimination). The significance level was set at $p < .05$ and Greenhouse-Geisser corrected when appropriate. When required, post-hoc analyses were conducted with Newman-Keuls tests.

3. Results

3.1. Perceptual scores

The main effect of task was significant ($F(3,27) = 4.34, p = .05$ – see Fig. 2), with a lower percentage of correct responses in the phoneme discrimination task than in the other tasks (on average, 98% (± 1) in the control task, 98% (± 2) in the phoneme identification task, 98% (± 2) in the syllable discrimination task and 89% (± 4) in

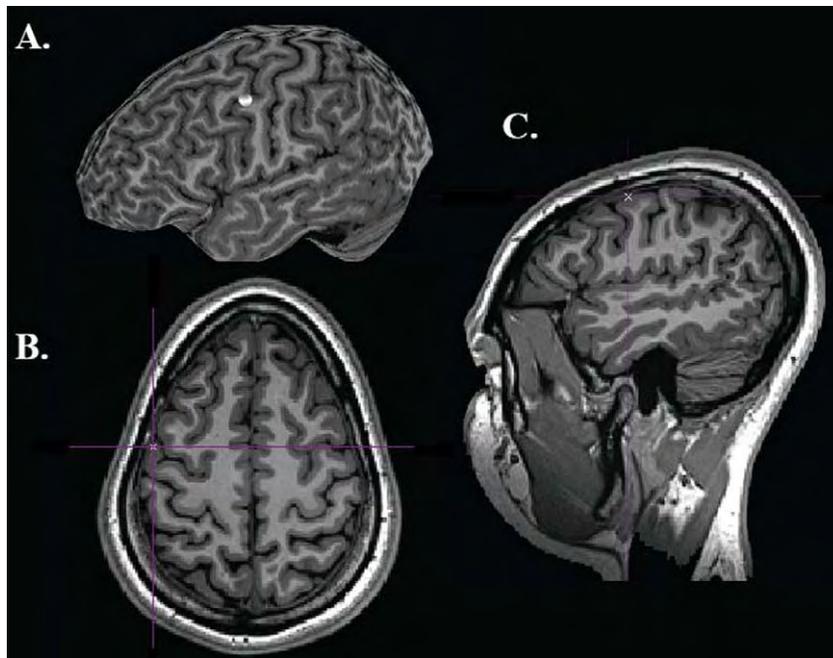


Fig. 1. Example of localization of the stimulation site in the superior portion of the left ventral premotor cortex from surface rendering (A) and axial (B) and sagittal (C) views taken from a structural MRI series of a participant.

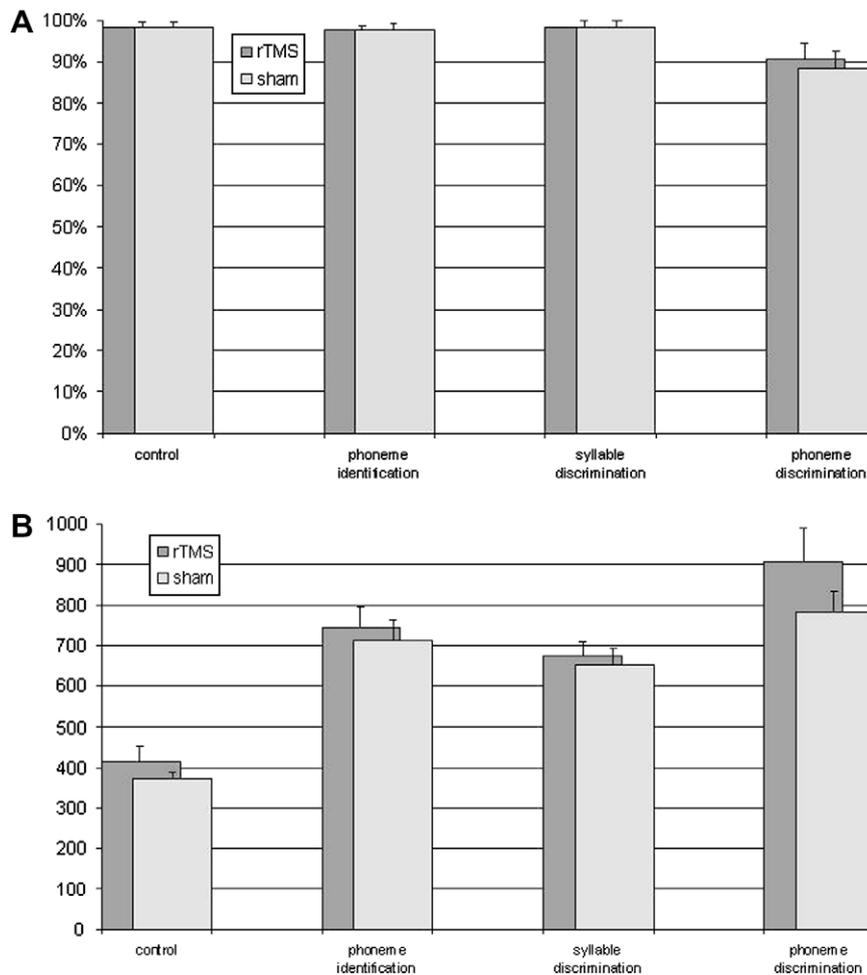


Fig. 2. Percentage of correct responses (A) and reaction times (B) for each task after rTMS and sham stimulation. Error bars represent standard errors of the mean.

the phoneme discrimination task; all p 's < .03). Neither the stimulation mode ($F(1,9) = 0.56$) nor the interaction between the two variables ($F(3,27) = 0.20$) were significant.

3.2. Reaction times

Significant differences were observed across tasks ($F(3,27) = 43.31$, $p < .00001$), with faster RTs in the control task and slower RTs in the phoneme discrimination task as compared to the other tasks (on average, 395 ms (± 27) in the control task, 730 ms (± 49) in the phoneme identification task, 663 ms (± 39) in the syllable discrimination task and 846 ms (± 65) in the phoneme discrimination task; all p 's < .01). The main effect of stimulation mode was also significant ($F(1,9) = 14.62$, $p < .005$), with slower RTs after rTMS than after sham stimulation (on average, 686 ms (± 51) vs. 631 ms (± 39)). Of main interest is the significant interaction between the two variables ($F(3,27) = 2.84$, $p = .05$), with significantly slower RTs after rTMS than after sham stimulation only in the phoneme discrimination task (909 ms (± 80) vs 783 ms (± 50); $p < .001$; all other comparisons not significant).

4. Discussion

Compared to sham stimulation, low-frequency rTMS applied over the svPMC had an effect on the response latencies in the phoneme discrimination task, which resulted in slower RTs without affecting the accuracy of the response. No specific effect caused by identical stimulation over the same cortical region was observed in the phoneme identification and the syllable discrimination tasks, nor in the visual matching task. Before we discuss these results, it is important to consider some inherent limitations of the TMS technique, related to the duration of post-stimulation effects and inter-participant anatomical differences (Federoes, Tononi, & Postle, 2007; Fitzgerald et al., 2006). The stimulation parameters for the present study were selected according to a recent and comprehensive review of the literature on the effects of stimulation of both motor and premotor cortex (Fitzgerald et al., 2006). Notably, the selected stimulation parameters have been shown repeatedly to produce inhibitory modulation of motor cortical excitability lasting up to 10 min. As a result, the duration of each experimental session was kept to about 6 min. Furthermore, the stimulation site was localized for all participants using frameless stereotaxy and individual MRI-to-head co-registration. Although we did not perform fMRI in order to individually determine the precise part of the left svPMC involved in speech perception, the stimulation site was the same as reported by Wilson et al. (2004), and was very close to those reported in two other fMRI studies (Meister et al., 2007; Pulvermuller et al., 2006). It is also worthwhile noting that the peak coordinates within the svPMC reported by Meister and colleagues (2007, p. 1693) was found to be consistent across their fifteen participants (mean MNI coordinates in mm \pm SD: -53 ± 4 and -4 ± 4 and 49 ± 2), the maximum x , y , and z distances between the peak coordinates (14, 12, 8 mm, respectively) being compatible with the spatial resolution of TMS (Walsh & Cowey, 2000). Finally, it is also likely that the rTMS effect observed in the phoneme discrimination task was not due to nonspecific attentional and/or sensory effects induced by the TMS pulses, because no such effect was observed for the other auditory tasks nor during the visual matching task.

In spite of these possible limitations, the specific increase of RT in the phoneme discrimination task, and the absence of any effect in the phoneme identification and syllable discrimination tasks, are in line with previous brain imaging, rTMS and electrocortical mapping studies of speech perception and phonological processing. Previous brain imaging studies on auditory phonological process-

ing have consistently reported left-sided activations of the posterior (opercular) part of the inferior frontal gyrus and of the ventral premotor cortex, in addition to temporal auditory regions (for a review, see Démonet, Thierry, & Cardebat, 2005; Poldrack et al., 1999; Vigneau et al., 2006). These studies typically involved phoneme monitoring and phoneme discrimination tasks in which participants had to separate a phoneme from a continuous acoustic signal in order to compare speech segments. For instance, Zatorre and colleagues (1992, Zatorre, Meyer, Gjedde, and Evans, 1996) and Burton et al. (2000) demonstrated increased frontal activity near the border of the left inferior frontal gyrus and the premotor cortex during a phoneme discrimination task, wherein participants were required to discriminate initial or final consonants in pairs of CVC syllables (e.g., /fat/-/tid/, /dip/-/ten/), compared to a pitch or a tone discrimination task. Given the involvement of these regions in the planning and execution of speech gestures (Bohland & Guenther, 2006; Gracco, Tremblay, & Pike, 2005; Sörös et al., 2006), it has been suggested that auditory-motor recoding and articulatory-based representations are recruited during phonological processing depending on the use of phonemic segmentation and working memory demands (Démonet, Price, Wise, & Frackowiak, 1994; Démonet et al., 1992; Zatorre et al., 1992, 1996; Burton & Small, 2006; Burton et al., 2000; Hickok & Poeppel, 2000; Hickok & Poeppel, 2004; Hickok & Poeppel, 2007). Our results are consistent with this interpretation and provide direct evidence for the functional participation of the svPMC in segmenting the speech stream into constituent phonemes.

In addition to results from svPMC, it has also been shown that a temporary disruption of the activity in the left posterior inferior frontal gyrus by means of rTMS disrupts participants' ability to perform phoneme discrimination (Romero et al., 2006). In that study, participants were required to make same/different judgments on visually presented word pairs, beginning with the same grapheme but which could be translated into one of two possible sounds. Compared to a baseline without TMS and rTMS over the vertex, rTMS over the left inferior frontal gyrus increased RTs and decreased accuracy. Increased response latency after rTMS of the left posterior inferior frontal cortex was also observed when participants had to decide whether two visually presented words rhymed or not (Gough, Nobre, & Devlin, 2005). The similarity of our rTMS results and those of Romero et al. (2006) and Gough et al. (2005) involving visually presented words suggests that the inferior frontal gyrus and the svPMC are both recruited for speech processing under conditions of increased working memory demands and articulatory rehearsal. Moreover, both areas appear to contribute in a top-down or feedforward manner to influence temporal lobe processing when phonemic segmentation is needed.

Despite this evidence, electrocortical mapping studies in epilepsy surgery patients (Boatman, 2004; Boatman & Miglioretti, 2005) and our results for simple phonetic and syllable discrimination indicate that stimulating the left inferior frontal lobe or premotor cortex does not impair participants' ability to discriminate CV syllables pairs in the absence of a sensory challenge or increased task demands. Both phoneme identification and syllable discrimination most likely rely solely on auditory analysis especially for syllable onset stop consonant voicing (Boatman, 2004; Burton et al., 2000). These processes are not influenced by inhibitory stimulation of the left premotor cortex or inferior frontal gyrus. In contrast, the phoneme discrimination task required a more complex analysis and segmentation of the sounds in the sequence (Burton & Small, 2006; Burton et al., 2000) and the disruptive rTMS effect strongly suggests that the left svPMC plays a functional role in speech segmentation recruited with increased task demands especially working memory.

The present results argue for a tight connection between speech perception and production systems (Aboitiz & Garcia, 1997; Aboi-

tiz, Garcia, Bosman, & Brunetti, 2006; Arbib, 2005; Callan, Jones, Callan, & Akahane-Yamada, 2004; Hickok & Poeppel, 2000, 2004, 2007; Rizzolatti & Arbib, 1998; Skipper et al., 2007; Wilson & Iacoboni, 2006). Notably, these results are fully consistent with the dual-stream model (Hickok and Poeppel, 2001, 2004, 2007) and the recruitment of a dorsal auditory-motor circuit involved in mapping sound onto articulatory-based representations. In this model, it is proposed that early cortical stages of speech processing involve auditory fields in the superior temporal gyrus. The ventral stream then projects to frontal motor regions (i.e., the posterior part of the inferior frontal gyrus and the premotor cortex), the primary function of which is to serve speech development and the acquisition of a new vocabulary. When a child learns to articulate speech sounds, it may provide a mechanism by which sensory representations of speech can be stored and compared against articulatory production. This comparison would be used to shape future productions. Importantly, the dorsal stream is not considered to be a critical component of speech perception in adults under normal listening conditions. Indeed, motor representations of speech are thought to be used strategically to assist in working memory and sub-lexical task performance, whenever translation of phonological information to an articulatory code is required. Also consistent with this view, the absence of any TMS-related modulation of participants' performance in both the phoneme identification and the syllable discrimination tasks does not fit with an obligatory functional role of the motor system in speech perception. From this view, our results cannot directly account for some neurobiological models (Callan et al., 2004; Skipper et al., 2005; Skipper et al., 2007; Wilson & Iacoboni, 2006) which postulate, contrary to Hickok and Poeppel's dual-stream model, that speech motor centers mediate speech perception under normal listening conditions, by constraining phonetic interpretation of the incoming sensory information. Note, however, that these models also assume that the speech motor centers are strongly recruited depending on the ambiguity of the sensory inputs, that is when the mapping between sensory information and phonetic categories is not sufficiently deterministic. In light of the present results, this latter proposal, recently confirmed by an rTMS study (Meister et al., 2007), and indirectly supported by fMRI studies showing an increased activation of the speech motor centers during the auditory identification of non-native versus native phonemes (e.g., Callan et al., 2004; Wilson & Iacoboni, 2006), during intelligible versus masked or distorted speech (e.g., Binder, Liebenthal, Possing, Medler, & DouglasWard, 2004; Zekveld, Heslenfeld, Festen, & Schoonhoven, 2006), and during the audiovisual observation of phonetically conflicting compared to matching vowels/syllables (e.g., Jones & Callan, 2003; Ojanen et al., 2005; Pekkola et al., 2006; Skipper et al., 2007) can be refined. It appears that the motor system reacts to noise or novelty or mismatch by enhancing the auditory signal to resolve signal ambiguity. In addition, when the perceptual task requires additional signal processing, the motor system is recruited to facilitate the processing requirements. From this view, what could be the function of the motor activity observed during passive speech perception (Fadiga et al., 2002; Ojanen et al., 2005; Watkins & Paus, 2004; Watkins et al., 2003, 2004; Pekkola et al., 2006; Pulvermuller et al., 2006; Roy et al., 2008; Skipper et al., 2005; Skipper et al., 2007; Wilson & Iacoboni, 2006)? One possibility comes from task complexity, participants being highly accurate in all three tasks with ceiling effects observed in both the phoneme identification and the syllable discrimination tasks. We cannot rule out the possibility that ceiling effects might have prevented possible TMS-related modulation of participants' accuracy to occur. Another possibility is that the motor system does not play a critical role in speech understanding under normal listening conditions and in the face of simple speech perceptual demands. In that case, the involvement of the motor sys-

tem would not be strictly intrinsic to speech comprehension but may facilitate conversational exchange by contributing to a common perceptuo-motor framework between speakers. In that case, speech motor resonance may represent a dynamic sensorimotor adaptation under the influence of the other talker's speech patterns, and in return may facilitate conversational interactions through convergent behaviors (Pardo, 2006; Sancier & Fowler, 1997).

By demonstrating a mediating role of the svPMC in the phoneme discrimination task, the present results argue against the view that speech perception relies exclusively on the auditory system without any role of the motor system, as postulated in purely auditory approaches of speech perception (for a review, see Diehl et al., 2004). Conversely, the absence of any TMS-related modulation of participants' performance in both the phoneme identification and the syllable discrimination tasks also challenges the view that perceiving speech is solely mediated by an articulatory code, as claimed in the motor theory of speech perception (Liberman & Mattingly, 1985; Liberman & Whalen, 2000; Liberman et al., 1967). Speech perception is best conceptualized as an interactive neural process involving reciprocal connections between sensory and motor areas whose connection strengths vary as a function of the perceptual task and the external environment. The observed disruptive rTMS effect suggests that the left svPMC plays a functional role in speech segmentation and is recruited with increased task demands under normal listening conditions.

Acknowledgments

We thank Jean-Luc Schwartz and Jérôme Aubin for their help. This study was supported by research grants from NSERC and CIHR (Canada) to V.L.G. and CNRS (France) to M.S., and by a Richard H. Tomlinson Research Fellowship and a travel grant from the 'PPP: Interactions multimodales' to M.S.

References

- Aboitiz, F., & Garcia, V. (1997). The evolutionary origin of the language areas in the human brain. A neuroanatomical perspective. *Brain Research Reviews*, 25, 381–396.
- Aboitiz, F., Garcia, V., Bosman, C., & Brunetti, E. (2006). Cortical memory mechanisms and language origins. *Brain and Language*, 98(1), 40–56.
- Arbib, M. A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, 28(2), 105–124.
- Binder, J. R., Liebenthal, E., Possing, E. T., Medler, D. A., & DouglasWard, B. (2004). Neural correlates of sensory and decision processes in auditory object identification. *Nature Neuroscience*, 7, 295–301.
- Boatman, D. F. (2004). Cortical bases of speech perception: Evidence from functional lesion studies. *Cognition*, 92, 47–65.
- Boatman, D. F., & Miglioretti, D. L. (2005). Cortical sites critical for speech discrimination in normal and impaired listeners. *Journal of Neuroscience*, 25(23), 5475–5480.
- Bohland, J. W., & Guenther, F. H. (2006). An fMRI investigation of syllable sequence production. *Neuroimage*, 32(2), 821–841.
- Burton, M. W., & Small, S. L. (2006). Functional neuroanatomy of segmenting speech and nonspeech. *Cortex*, 42, 644–651.
- Burton, M. W., Small, S. L., & Blumstein, S. E. (2000). The role of segmentation in phonological processing: An fMRI investigation. *Journal of Cognitive Neuroscience*, 12(4), 679–690.
- Callan, D. E., Jones, J., Callan, A., & Akahane-Yamada, R. (2004). Phonetic perceptual identification by native- and second-language speakers differentially activates brain regions involved with acoustic phonetic processing and those involved with articulatory-auditory/orosensory internal models. *Neuroimage*, 22, 1182–1194.
- Démonet, J.-F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J.-L., Wise, R., et al. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain*, 115, 1753–1768.
- Démonet, J.-F., Price, C., Wise, R., & Frackowiak, R. S. J. (1994). A pet study of cognitive strategies in normal subjects during language tasks: Influence on phonetic ambiguity and sequence processing on phoneme monitoring. *Brain*, 117(4), 671–682.
- Démonet, J. F., Thierry, G., & Cardebat, D. (2005). Renewal of the neurophysiology of language: Functional neuroimaging. *Physiological Reviews*, 85(1), 49–95.

- Devlin, J. T., & Watkins, K. E. (2007). Stimulating language: Insights from TMS. *Brain*, 130, 610–622.
- Diehl, R. L., Lotto, A. J., & Holt, L. L. (2004). Speech perception. *Annual Review of Psychology*, 55, 149–179.
- Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience*, 15, 399–402.
- Federoes, E., TONI, G., & Postle, B. S. (2007). Direct evidence for a prefrontal contribution to the control of proactive interference in verbal working memory. *Proceedings of the National Academy of Sciences of the United States of America*, 103(51), 19530–19534.
- Fitzgerald, P. B., Fountain, S., & Daskalakis, Z. J. (2006). A comprehensive review of the effects of rTMS on motor cortical excitability and inhibition. *Clinical Neurophysiology*, 117(12), 2584–2596.
- Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, 13(3), 361–377.
- Gough, P. M., Nobre, A. C., & Devlin, J. T. (2005). Dissociating linguistic processes in the left inferior frontal cortex with transcranial magnetic stimulation. *Journal of Neuroscience*, 25, 8010–8016.
- Gracco, V. L., Tremblay, P., & Pike, G. B. (2005). Imaging speech production using fMRI. *Neuroimage*, 26, 294–301.
- Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Science*, 4(4), 131–138.
- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, 92, 67–99.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8, 393–402.
- Jones, J., & Callan, D. E. (2003). Brain activity during audiovisual speech perception: An fMRI study of the McGurk effect. *Neuroreport*, 14(8), 1129–1133.
- Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74, 431–461.
- Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21, 1–36.
- Liberman, A. M., & Whalen, D. H. (2000). On the relation of speech to language. *Trends in Cognitive Science*, 3(7), 254–264.
- Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D., & Iacoboni, M. (2007). The essential role of premotor cortex in speech perception. *Current Biology*, 17(19), 1692–1696.
- Ojanen, V., Möttönen, R., Pekkola, J., Jääskeläinen, I. P., Joensuu, R., Autti, T., et al. (2005). Processing of audiovisual speech in Broca's area. *Neuroimage*, 25, 333–338.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97–114.
- Pardo, J. (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America*, 119(4), 2382–2393.
- Pascual-Leone, A., Walsh, V., & Rothwell, J. (2000). Transcranial magnetic stimulation in cognitive neuroscience: Virtual lesion, chronometry, and functional connectivity. *Current Opinion in Neurobiology*, 10, 232–237.
- Pekkola, J., Laasonen, M., Ojanen, V., Autti, T., Jaaskelainen, L. P., & Kujala, T. (2006). Perception of matching and conflicting audiovisual speech in dyslexic and fluent readers: An fMRI study at 3T. *Neuroimage*, 29(3), 797–807.
- Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage*, 10, 15–35.
- Pulvermuller, F., Huss, M., Kherif, F., Moscoso del Prado Martin, F., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences of the United States of America*, 103(20), 7865–7870.
- Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, 21, 188–194.
- Romero, L., Walsh, V., & Papagno, C. (2006). The neural correlates of phonological short-term memory: A repetitive transcranial magnetic stimulation study. *Journal of Cognitive Neuroscience*, 18(7), 1147–1155.
- Rossini, P. M., Barker, A. T., Berardelli, A., Caramia, M. D., Caruso, G., Cracco, R. Q., et al. (1994). Non-invasive electrical and magnetic stimulation of the brain, spinal cord and roots: Basic principles and procedures for routine clinical application. Report of an IFCN committee. *Electroencephalography and Clinical Neurophysiology*, 91(2), 79–92.
- Roy, A. C., Craighero, L., Fabbri-Destro, M., & Fadiga, L. (2008). Phonological and lexical motor facilitation during speech listening: A transcranial magnetic stimulation study. *Journal of Physiology-Paris*, 102(1–3), 101–105.
- Sancier, M., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*, 25, 421–436.
- Schwartz, J.-L., Abry, C., Boë, L.-J., & Cathiard, M. A. (2002). Phonology in a theory of perception-for-action-control. In J. Durand & B. Lacks (Eds.), *Phonology: From phonetics to cognition* (pp. 240–280). Oxford: Oxford University Press.
- Skipper, J. I., Nusbaum, H. C., & Small, S. L. (2005). Listening to talking faces: Motor cortical activation during speech perception. *Neuroimage*, 25, 76–89.
- Skipper, J. I., Van Wassenhove, V., Nusbaum, H. C., & Small, S. L. (2007). Hearing lips and seeing voices: How cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex*, 17(10), 2387–2399.
- Sörös, P., Sokoloff, L. G., Bose, A., McIntosh, A. R., Graham, S. J., & Stuss, D. T. (2006). Clustered functional MRI of overt speech production. *Neuroimage*, 32(1), 376–387.
- Sundara, M., Namasivayam, A. K., & Chen, R. (2001). Observation–execution matching system for speech: A magnetic stimulation study. *Neuroreport*, 12(7), 1341–1344.
- Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., et al. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *Neuroimage*, 30(4), 1414–1432.
- Walsh, V., & Cowey, A. (2000). Transcranial magnetic stimulation and cognitive neuroscience. *Nature Reviews Neuroscience*, 1, 73–79.
- Wassermann, E. M. (1998). Risk and safety of repetitive transcranial magnetic stimulation. Report and suggested guidelines from the international workshop on the safety of repetitive transcranial magnetic stimulation, June 5–7, 1996. *Electroencephalography and Clinical Neurophysiology*, 108(1), 1–16.
- Watkins, K. E., & Paus, T. (2004). Modulation of motor excitability during speech perception: The role of Broca's area. *Journal of Cognitive Neuroscience*, 16(6), 978–987.
- Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, 41(3), 989–994.
- Wilson, S. M., & Iacoboni, M. (2006). Neural responses to non-native phonemes varying in producibility: Evidence for the sensorimotor nature of speech perception. *Neuroimage*, 33(1), 316–325.
- Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, 7, 701–702.
- Zatorre, R., Evans, A., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science*, 256, 846–849.
- Zatorre, R. J., Meyer, E., Gjedde, A., & Evans, A. C. (1996). PET studies of phonetic processing of speech: Review, replication, and reanalysis. *Cerebral Cortex*, 6(1), 21–30.
- Zekveld, A. A., Heslenfeld, D. J., Festen, J. M., & Schoonhoven, R. (2006). Top-down and bottom-up processes in speech comprehension. *Neuroimage*, 32, 1826–1836.