Available online at www.sciencedirect.com
SciVerse ScienceDirect
www.elsevier.com/locate/brainres

Research Report

1725

A mediating role of the auditory dorsal pathway in selective adaptation to speech: A state-dependent transcranial magnetic stimulation study ☆

Krystyna Grabski[a,c,*], *Pascale Tremblay*[b], *Vincent L. Gracco*[c,d,e], *Laurent Girin*[a], *Marc Sato*[a,c,*]

[a]*GIPSA-lab, Département Parole & Cognition, CNRS & Grenoble Université, France*
[b]*Université Laval, Département de Réhabilitation, Centre de Recherche de l'Institut Universitaire en Santé Mentale de Québec, Québec, Canada*
[c]*Centre for Research on Brain, Language and Music, McGill University, Montreal, Canada*
[d]*School of Communication Sciences & Disorders, McGill University, Montreal, Canada*
[e]*Haskins Laboratories, New Haven, USA*

ARTICLE INFO

Article history:
Accepted 21 March 2013
Available online 29 March 2013

Keywords:
Speech perception and speech production
Auditory dorsal pathway
Sensorimotor interaction
Selective adaptation
Transcranial magnetic stimulation

ABSTRACT

In addition to sensory processing, recent neurobiological models of speech perception postulate the existence of a left auditory dorsal processing stream, linking auditory speech representations in the auditory cortex with articulatory representations in the motor system, through sensorimotor interaction interfaced in the supramarginal gyrus and/or the posterior part of the superior temporal gyrus. The present state-dependent transcranial magnetic stimulation study is aimed at determining whether speech recognition is indeed mediated by the auditory dorsal pathway, by examining the causal contribution of the left ventral premotor cortex, supramarginal gyrus and posterior part of the superior temporal gyrus during an auditory syllable identification/categorization task. To this aim, participants listened to a sequence of /ba/ syllables before undergoing a two forced-choice auditory syllable decision task on ambiguous syllables (ranging in the categorical boundary between /ba/ and /da/). Consistent with previous studies on selective adaptation to speech, following adaptation to /ba/, participants responses were biased towards /da/. In contrast, in a control condition without prior auditory adaptation no such bias was observed. Crucially, compared to the results observed without stimulation, single-pulse transcranial magnetic stimulation delivered at the onset of each target stimulus interacted with the initial state of each of the stimulated brain area by enhancing the adaptation effect. These results demonstrate that the auditory dorsal pathway contribute to auditory speech adaptation.

☆*Author contributions*: Conceived and designed the experiments: KG PT VLG LG MS. Performed the experiments: KG PT. Analyzed the data: KG MS. Wrote the paper: KG PT VLG MS.

*Correspondence to: GIPSA-LAB, UMR CNRS 5216, Grenoble Université, 1180, avenue centrale, BP 25, 38040 Grenoble Cedex 9, France. Fax: +33 476 824 335.
E-mail addresses: krystyna.m.grabski@gmail.com (K. Grabski), marc.sato@gipsa-lab.inpg.fr (M. Sato).

0006-8993/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.brainres.2013.03.024

1. Introduction

How do listeners process the speech signal to recover phonetic information? Based on a constructivist approach of perception (Helmholtz, 1867), several theories of speech perception argue that the speaker and the listener share a common repertoire of sensory and motor primitives and that speech perception is partly driven by a process of internal sensory-motor simulation that serves to constrain the phonetic interpretation of the sensory inputs (Liberman et al., 1967; Stevens and Halle, 1967; Liberman and Mattingly, 1985; Liberman and Whalen, 2000; Schwartz et al., 2002, 2012; for recent reviews, Galantucci et al., 2006; Schwartz et al., 2008). The constructivist approach is also at the core of recent distributed brain network models of speech processing (Callan et al., 2004; Wilson and Iacoboni, 2006; Hickok and Poeppel, 2007; Skipper et al., 2007; Poeppel et al., 2008; Rauschecker and Scott, 2009; Hickok et al., 2011; Rauschecker, 2011). These models postulate the existence of a left posterior dorsal ('how') processing stream, linking auditory speech representations in the auditory cortex and articulatory representations in the ventral premotor cortex (vPM) and the posterior part of the inferior frontal gyrus (pIFG), with sensorimotor interaction converging in the supramarginal gyrus (SMG; Rauschecker and Scott, 2009; Rauschecker, 2011) or in area SPT (a brain region within the planum temporale near the parieto-temporal junction; Hickok and Poeppel, 2007). This auditory dorsal pathway allows internal sensory-to-motor and motor-to-sensory projection/mapping between auditory and motor representations, the role of which is to provide predictive coding schemes to compare top-down motor predictions with bottom-up sensory information to recover phonetic interpretation. In support of these models, motor activity during auditory, visual and auditory-visual speech perception has been observed in a considerable number of brain imaging and neurophysiological studies (e.g., Sundara et al., 2001; Fadiga et al., 2002; Nishitani and Hari, 2002; Calvert and Campbell, 2003; Callan et al., 2003; Paulesu et al., 2003; Watkins et al., 2003; Callan et al., 2004; Watkins and Paus, 2004; Wilson et al., 2004; Ojanen et al., 2005; Pekkola et al., 2006; Skipper et al., 2005; Pulvermuller et al., 2006; Wilson and Iacoboni, 2006; Skipper et al., 2007; Roy et al., 2008; Callan et al., 2010; Sato et al., 2010; Tremblay and Small, 2011; Alho et al., 2012; Grabski et al., 2013). However, for all their importance, these results are intrinsically correlational and cannot address causality. Weak perturbations in auditory syllable identification tasks have been obtained by temporarily disrupting the activity of components of the cortical motor system in the presence of masking noise (Meister et al., 2007; d'Ausilio et al., 2009, 2012; but see Sato et al., 2011 for postperceptual bias effects), when processing acoustically ambiguous syllables (Möttonen and Watkins, 2009) or performing a task requiring phonemic segmentation or loading on working memory (Boatman, 2004; Gough et al., 2005; Romero et al., 2006; Sato et al., 2009).

In keeping with these results, the goal of the present study was to further investigate the contribution of the auditory dorsal pathway, specifically the left vPM, SMG and pSTG, to speech perception. A state-dependent TMS paradigm (e.g., Silvanto et al., 2008; Cattaneo and Silvanto, 2008; Cattaneo et al., 2010; Cattaneo et al., 2010) was used to examine the causal contribution of these brain regions in an auditory syllable identification/categorization task. This paradigm is based on sensory-induced perceptual adaptation prior to online stimulation and on the assumption that the effects of online TMS depend on the initial state of the stimulated brain area. More specifically, perceptual adaptation is first induced by repetitively presenting a stimulus, without stimulation, with the goal of tuning specific neural populations in the targeted brain area that code particular features of the stimulus. In a subsequent identification/categorization task, if an online single-pulse TMS then interacts with this initial state manipulation (for example by facilitating detection of a specific feature of the repeated stimulus), this likely indicates that neurons in the targeted brain region were indeed tuned to the adapting stimulus. A selective adaptation paradigm was used to induce a categorical perceptual shift on subsequent target syllables. Selective adaptation to speech refers to the repeated presentation of a particular speech stimulus that causes a reduction in the frequency with which that stimulus is reported in subsequent identification trials. For example, in the seminal study by Eimas and Corbit (1973), listeners had to categorize syllables from a/ba/–/p^{h}a/ continuum. Listening to repeated presentation of /ba/ syllables prior to the identification/categorization task induced fewer perceived /ba/ than /p^{h}a/ syllables (at the categorical boundary) than observed without adaptation, while the reverse was true when /p^{h}a/ was the repeated stimulus. Using the same paradigm, Cooper and Lauritsen (1974) further demonstrated that prolonged listening to a syllable with an initial voiceless stop consonant caused subject to produce a shorter voice onset time for the same syllable in a subsequent production task. Since both perceptual and motor adaptive changes occur after prolonged listening to a speech sound, selective adaptation to speech is likely to mediate both speech perception and production through fatigue of specialized phonetic feature detectors and/or criterion-setting operations (for reviews, see Cooper, 1979; Samuel, 1986). In the present study, participants listened to a sequence of /ba/ syllables before undergoing a two forced-choice auditory syllable decision task on ambiguous syllables (ranging in the categorical boundary between /ba/ and /da/). Based on previous studies on selective adaptation, auditory adaptation should induce short-term perceptuo-motor changes and bias participants' responses towards /da/ in the subsequent syllable decision task, as compared to a control condition without prior auditory adaptation. In additional conditions, single-pulse TMS over the left vPM, SMG or pSTG were delivered at the onset of each target stimulus using frameless stereotaxy and individual MRI-to-head co-registration. Compared to a control condition (SHAM) performed without any stimulation, contrastive results on selective adaptation observed after TMS applied over the left vPM, SMG and/or pSTG should demonstrate whether these regions causally contribute to speech categorization/recognition.

2. Results

In order to individually determine the three syllables (c^-, c^0, c^+) ranging in the categorical boundary between/ba/ and/da/, participants first underwent a categorical perception task,

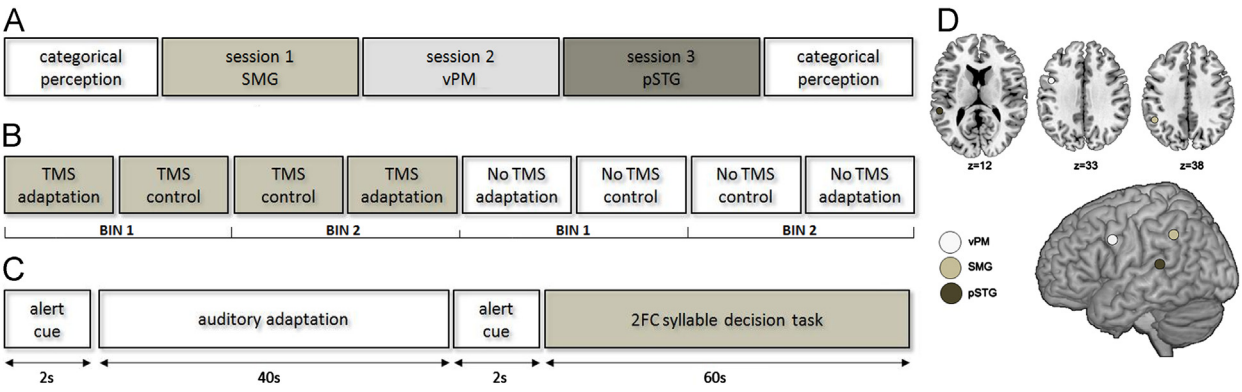

Fig. 1 – Experimental design. (A) The experiment consisted of a categorical perception task (without TMS), three TMS sessions (according to each stimulation site: pSTG, SMG, vPM) consisting of a 2FC syllable decision task, and a final categorical perception task (without TMS). (B) With the exception of the stimulation site, the experimental procedure was identical in the three TMS sessions. The TMS and SHAM conditions each occurred in four successive experimental blocks, with the auditory adaptation and control conditions occurring once in both the first and last two blocks (BIN1, BIN2). (C) Each block consisted of an adaptation or a control task and a subsequent 2FC syllable decision task with or without stimulation. The order of the stimulation site, the stimulation mode and the adaptation mode were counterbalanced across participants. (D) Mean localization of pSTG, SMG and vPM stimulation sites for all participants rendered on axial slices and cortical surface on a standard brain template using the MRICRON software (http://www.sph.sc.edu/comd/rorden/mricron/).

without TMS. Three two forced-choice syllable decision tasks were then performed on these syllables using TMS on each stimulation site (see Fig. 1 and Experimental procedures). Finally, a second categorical perception task, without TMS, was performed.

2.1. *Categorical decision tasks*

Fig. 2 displays the mean proportion of /ba/ responses to the eleventh stimuli of the /ba/-/da/ auditory continuum (c^{-5} to c^{+5}) in the two categorical decision tasks (performed without stimulation, but in the presence of TMS acoustic noise, before and after the TMS experiment). As expected, the proportion of /ba/ responses increased significantly as the auditory stimulus moved from /da/ to /ba/ ($F_{(10,110)}=137.7$, $p<.001$). However, no modulation of categorical perception was observed before and after TMS, with neither the effect of experimental session or the 'stimulus x session' interaction being significant. This result indicates that selective adaptation to speech observed in the TMS experiment (see below) did not provide long-lasting after-effects, with no perceptual shift occurring in the final categorical decision task.

2.2. *TMS-2FC syllable decision tasks*

Fig. 3 displays the mean proportion of/ba/ responses and median RTs according to the stimulation mode (TMS vs. SHAM), the stimulation site (pSTG, SMG, vPM) and the adaptation mode (ADAPTATION, CONTROL). Fig. 4 displays the strength of the adaption effect (proportion of /ba/ responses without prior auditory adaptation subtracted from those with prior adaptation) according to the stimulation mode (TMS vs. SHAM) and the stimulation site (pSTG, SMG, vPM).

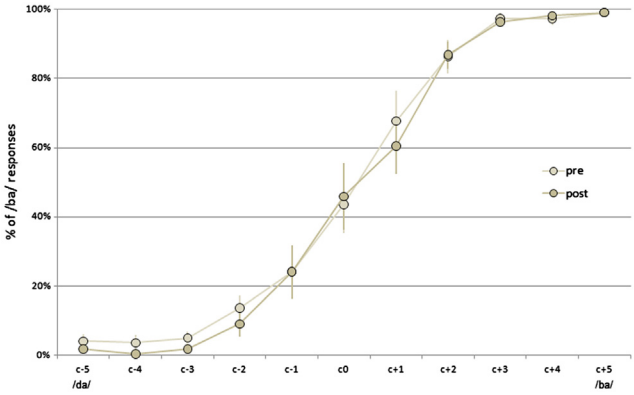

Fig. 2 – Proportion of /ba/ responses as a function of the eleven stimuli of the /ba/–/da/ auditory continuum (c^{-5} to c^{+5}) in the two categorical decision tasks (before and after the TMS experiment). Error bars represent standard errors of the mean (SEM).

2.2.1. *Reaction times*

Overall, the mean RT was 641 ms (±30); it was similar in the three TMS sessions (on average, 621 ms (±26), 660 ms (±37) and 642 ms (±34) for SMG, pSTG and vPM, respectively; see Fig. 3, bottom). The 'stimulation site × stimulation mode' interaction was the only significant effect ($F_{(1,9)}=5.29$, $p=.02$), with responses after TMS over SMG and pSTG being slower compared to SHAM (on average, −36 ms and −34 ms, respectively) while the inverse effect was observed after TMS over vPM (+28 ms). No other significant effect or interaction was observed.

2.2.2. *Perceptual scores*

Overall, the mean identification score was of 46% (±3); it was similar across the three TMS sessions (on average, 44% (±6), 48% (±3) and 47% (±4) of /ba/ responses for SMG, pSTG and

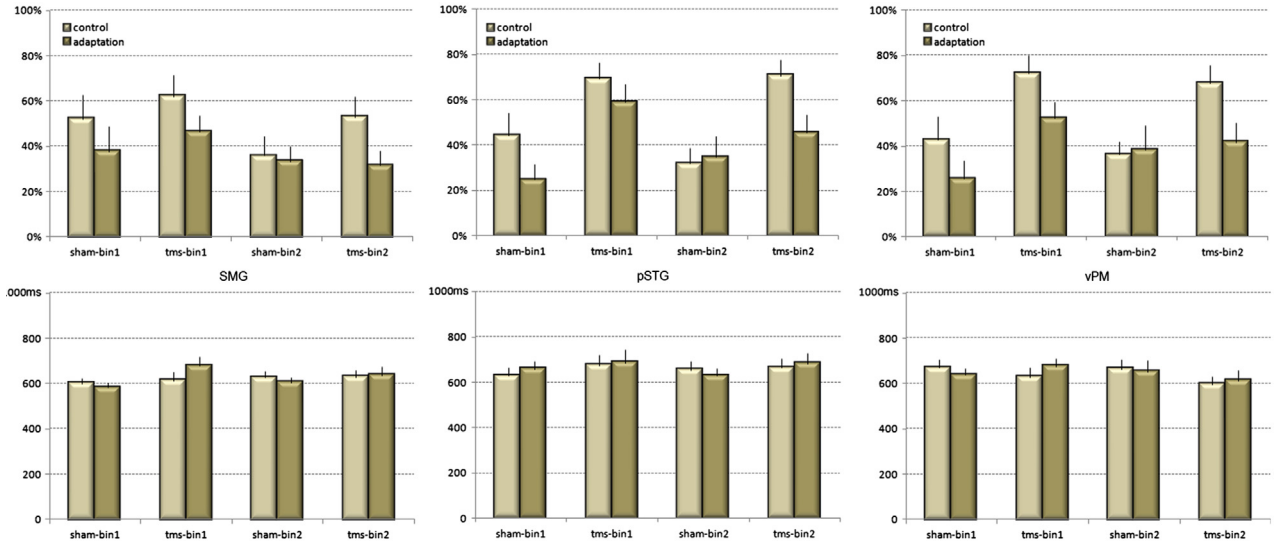

Fig. 3 – Proportion of /ba/ responses and median RTs observed in the syllable decision tasks according to the stimulation site (pSTG, SMG, vPM), the stimulation mode (TMS, SHAM), the adaptation mode (ADAPTATION, CONTROL) and the bin (BIN1, BIN2). Error bars represent standard errors of the mean.

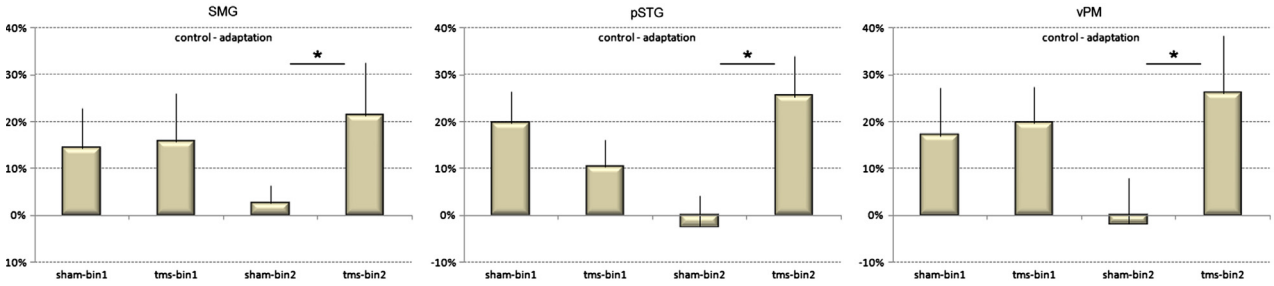

Fig. 4 – Size of the adaptation effect (proportion of /ba/ responses without prior auditory adaptation subtracted from those with prior adaptation) observed in the syllable decision tasks according to the stimulation site (pSTG, SMG, vPM), the stimulation mode (TMS, SHAM) and the bin (BIN1, BIN2). Error bars represent standard errors of the mean.

vPM, respectively; see Fig. 3, top). This confirmed that c^-, c_0 and c^+ stimuli located at the categorical boundary between /ba/ and/da/ in each TMS session.

A 4-WAY repeated-measure ANOVA showed a strong selective adaptation, as indicated by a significant main effect of the adaptation mode ($F_{(1,9)}=17.2$, $p=.003$). This selective adaption effect, which occurred in the three TMS sessions (vPM, SMG, pSTG) and the two stimulation modes, corresponded to the predicted decrease of /ba/ responses in the auditory adaptation condition compared to the control condition (on average, −9% and −19% for SMG/SHAM and SMG/TMS, −9% and −18% for pSTG/SHAM and pSTG/TMS, −8% and −23% for vPM/SHAM and vPM/TMS).

Second, a higher proportion of /ba/ responses was observed after TMS compared to SHAM ($F_{(1,9)}=8.4$, $p=.02$) with an increase of /ba/ responses after TMS occurring in the three TMS sessions and the two adaptation modes (on average, +14% and +4% for SMG/CONTROL and SMG/ADAPTATION, +32% and +23% for pSTG/CONTROL and pSTG/ADAPTATION, +30% and +15% for vPM/CONTROL and vPM/ADAPTATION). Moreover, this effect persisted from BIN1 to BIN2 with a non significant 'stimulation mode × bin' interaction ($F_{(1,9)}=.38$, $p=.55$; on average, +9% and +8% for SMG/BIN1 and SMG/BIN2, +30% and +25% for pSTG/BIN1 and pSTG/BIN2, +28% and +17% for vPM/BIN1 and vPM/BIN2). The most likely explanation of this effect is that the absence of tactile contact in the SHAM conditions might have potentially lowered the auditory (clicking) sound and, although the single pulse clicking sound occurred during the stimulus onset in both conditions (TMS, SHAM), bone-conducted sounds might have biased participants' responses towards /ba/ in the TMS conditions (Nikouline et al., 1999).

Crucially, a significant 'stimulation mode × adaptation mode × bin' interaction was observed ($F_{(1,9)}=6.5$, $p=.03$). Further analysis of the size of the adaptation effect showed that a selective adaptation was observed in BIN1 in both the TMS and SHAM conditions (on average, −14% and −16% of /ba/ responses for SMG/SHAM and SMG/TMS, −20% and −10% for pSTG/SHAM and pSTG/TMS, −17% and −20% for vPM/SHAM and vPM/TMS, no comparisons significant; see Fig. 4).

However, in BIN2, while the adaptation effect almost disappeared in the SHAM condition, it remained present in the TMS condition (on average, –3% and –21% of /ba/ responses for SMG/SHAM and SMG/TMS, 3% and –25% for pSTG/SHAM and pSTG/TMS, 2% and –26% for vPM/SHAM and vPM/TMS, all comparisons significant). It should be noted that this interaction is not likely due to TMS acoustic artefacts since (1) the higher proportion of /ba/ responses after TMS was equally observed in BIN1 and BIN2, (2) selective adaptation was independently computed for each stimulation site and each bin for the TMS and SHAM conditions (i.e., (SHAM-CONTROL minus SHAM-ADAPTATION) vs. (TMS-control minus TMS-ADAPTATION) and (3) that the stimulation and single pulse clicking sound equally occurred with or without prior adaptation in the TMS condition. Hence, these results suggest that selective adaptation persisted in BIN2 because of TMS while it disappeared without stimulation (in the sham condition). No other significant effect or interaction was observed.

3. Discussion

The present study addressed the question of whether the auditory dorsal pathway mediates speech categorization by examining the causal contribution of the left vPM, SMG and pSTG in an auditory syllable categorization task. To this aim, auditory perceptual adaptation was first behaviorally induced, with the goal of tuning specific neural populations in the targeted brain areas. As expected, in a subsequent auditory syllable decision task on ambiguous syllables, a strong selective adaption was observed as compared to a control condition without prior auditory adaptation. Crucially, compared to the results observed without stimulation, single-pulse TMS delivered at the onset of each target stimulus over the left vPM, SMG or SMG, using frameless stereotaxy and individual MRI-to-head co-registration, interacted with the initial state of the stimulated brain area by enhancing the adaptation effect. The findings suggest that these regions are part of a network that contributes to auditory speech adaptation and provide further evidence for a mediating role of the dorsal pathway in speech categorization.

The state-dependent TMS paradigm consisted of a combination of auditory-induced perceptual adaptation and online TMS delivered in a subsequent syllable categorization task. This paradigm, known to enhance the functional selectivity and resolution of TMS (for a review, see Silvanto et al., 2008), was used to determine a possible causal contribution of three target regions in the categorization/decision process and in relation to the adapted stimulus. In order to determine, individually for each participant, the best range for ambiguous stimuli in the categorical boundary between /ba/ and /da/, participants first performed a categorical decision task on a /ba/–/da/ continuum, without stimulation. Three syllables centered on the midpoint of the discrimination function then served as target stimuli in the next TMS sessions. For these three syllables across all conditions and sessions, the mean identification scores were always near chance level (on average, 44%, 48% and 47% of/ba/ responses for SMG, pSTG and vPM, respectively; see Fig. 3), thus confirming the ambiguity of these targets. During the experimental sessions, prolonged listening to a/ba/ syllable induced a strong selective adaptation to speech (Eimas and Corbit, 1973), which led to fewer /ba/ responses in the subsequent categorization of ambiguous target syllables than observed in a similar categorization task without prior auditory adaptation. These results were observed equally in the three experimental sessions (on average, –14%, –13% and –15% of /ba/ responses for SMG, pSTG and vPM, respectively) consistent with previous selective adaptation studies (for reviews, see Cooper, 1979; Samuel, 1986). In addition, no perceptual shift was observed between the initial and final categorization tasks on the /ba/–/da/ continuum performed without stimulation (see Fig. 2), indicating that the selective adaptation to speech observed in the TMS sessions did not induce long-lasting perceptual after-effects. In sum, these results confirm the validity of the experimental paradigm used to induce selective adaptation to speech.

It should be noted that a higher proportion of /ba/ responses was observed after TMS compared to SHAM in each of the three TMS sessions and for each of the two adaptation modes. Since the single pulse clicking sound occurred during the stimulus onset in both TMS and SHAM conditions regardless of the auditory adaptation mode, the most likely explanation for the higher proportion of /ba/ responses (see Fig. 3), is that the absence of tactile contact in the SHAM conditions might have potentially lowered the auditory (clicking) sound and, conversely, bone-conducted sounds from single-pulse TMS might have contributed to bias participants' responses. Auditory evoked potentials from the acoustical click are known to be greatly affected by the position of the coil, with greater amplitude observed with the coil pressed against the scalp than with the coil placed 2 cm above the head (Nikouline et al., 1999). In the present study, it is therefore possible that the coil click propagated by air and bone in the TMS condition might have partially masked the consonantal burst and higher frequency transitions of the target syllables (see Methods) and biased participants' responses. Importantly, this higher proportion of /ba/ responses after TMS was observed in each of the two experimental blocks (BIN1 and BIN2).

Crucially, compared to SHAM stimulation, single-pulse TMS delivered at the onset of each target stimulus interacted with the initial state of the stimulated brain area by enhancing the adaptation effect in the last experimental block. In the first experimental blocks (BIN1), a strong adaptation effect was equally observed in the three stimulations sites for both TMS and SHAM conditions (on average, –15%, –15%, –18% of /ba/ responses for SMG, pSTG and vPM, respectively; see Fig. 4). This result indicates that TMS has no additive/modulatory effect on selective adaptation in the first experimental block compared to no stimulation, possibly due to ceiling-effects on adaptive changes. However, in the last experimental blocks (BIN2), while the adaptation effect almost disappeared in the SHAM condition (on average, –3%, +3% and +2% of/ba/ responses for SMG, pSTG and vPM, respectively; see Fig. 4), it remained present in the TMS condition (on average, –21%, –25% and –26% of /ba/ responses for SMG, pSTG and vPM, respectively; see Fig. 4).

As previously mentioned, it is worth noting that (1) the higher proportion of/ba/ responses likely due to TMS acoustic artefacts was similar across the two experimental blocks (BIN1 and BIN2), (2) for each stimulation site and each bin, selective adaptation was independently computed for the TMS and SHAM conditions, and (3) the stimulation and single pulse clicking sound occurred with or without prior adaptation in the TMS condition. For all these reasons, TMS acoustic artefacts cannot satisfactorily explain the strong adaptation effect observed in the three stimulations sites for TMS but not for SHAM conditions in the last experimental block. It has been suggested that selective adaption operates on speech perception and production through fatigue of specialized phonetic feature detectors and/or criterion-setting operations (for reviews, see Cooper, 1979; Samuel, 1986). However, these mechanisms might not have operated similarly throughout the experiment, especially with respect to cumulative and previous participants' biased responses. From that view, without additional stimulation, selective adaptation would disappear in the second experimental block due to competition and/or bias/criterion-settings mechanisms between repeated presentations of /ba/ and previous participants' biased responses towards /da/. However, the adaptation effect observed in the second experimental block after TMS indicates that modulation of cortical excitability in SMG, pSTG and vPM interacted with these processes by maintaining enhancing biased responses. Furthermore, the fact that the bin effect was replicated 3 times with 3 sham conditions in each stimulation site argues for the reliability of this effect (see Fig. 4).

In sum, although this effect appears to be due to a complicated interaction between selective adaptation and TMS, the observed state-dependency of the TMS-induced effects suggests that the auditory dorsal pathway contributes to auditory syllabic adaptation. Importantly, the size of this effect was similar across the three stimulated brain areas. Although our results do not specify any distinctive role for each of the three stimulated areas in speech adaptation/categorization, they suggest that the three regions (vPM, SMG and pSTG) of the auditory dorsal pathway work in concert during the categorization of ambiguous speech stimuli. Such possible mutual influence of these dorsal stream regions is consistent with their anatomical connections (for a recent review, see Dick and Tremblay, 2012) and with recent distributed neurobiological network models of speech processing (Callan et al., 2004; Wilson and Iacoboni, 2006; Hickok and Poeppel, 2007; Skipper et al., 2007; Poeppel et al., 2008; Rauschecker and Scott, 2009; Hickok et al., 2011; Rauschecker, 2011), which postulate bidirectional interactions between these regions during the categorization of ambiguous speech stimuli. Finally, it is worthwhile noting that whether other brain areas, notably those of the ventral pathway (Hickok and Poeppel, 2007; Rauschecker and Scott, 2009; Rauschecker, 2011), may also be involved in selective adaptation to speech.

Our results appear in line with previous TMS-adaptation studies showing that TMS interacts with initial state changes induced by visual adaptation or priming (e.g., Silvanto et al., 2008; Cattaneo and Silvanto, 2008; Cattaneo et al., 2010; Cattaneo et al., 2010; Cattaneo, 2010). Interestingly, it has been shown that repeated exposure to visually presented hand/foot transitive actions modulates behavioral performance in a subsequent visual identification task of similar actions when single-pulse TMS was applied over the vPM, the inferior parietal lobule and the superior temporal sulcus, compared to no stimulation (Cattaneo et al., 2010; Cattaneo, 2010). The opposite perspective, that is, the influence of motor behavior on perception, has also been investigated, with visual after-effect induced by blindfolded repeated motor performance of an object-directed action modulated by TMS applied over the left vPM (Cattaneo et al., 2011; see also Glenberg et al., 2010). These results are generally consistent with those obtained in recent studies using an fMRI adaptation paradigm to investigate the neural substrates of action goal coding. Repetition suppression effects, corresponding to a decrease in blood oxygen level-dependent signal to repeated stimuli (Grill-Spector and Malach, 2001; Grill-Spector et al., 2006), were indeed observed in the left vPM, inferior parietal lobule and adjacent intraparietal sulcus during repeated observation of manual actions (Dinstein et al., 2007; Chong et al., 2008; Lingnau et al., 2009; Kilner et al., 2009) as well as during a cross-modal paradigm, with response suppression observed when manual actions were first observed and then executed and vice versa (Kilner et al., 2009). These results have been largely discussed in the context of the human mirror-neuron system and its possible role in action goal coding (for reviews, see Rizzolatti et al., 2001; Rizzolatti and Craighero, 2004).

Regarding speech perception, although our results do not specify any distinctive role of the motor system, they are in keeping with previous TMS studies showing that temporary disruption of the activity of the left vPM or primary motor cortex induce modulation of performance in auditory syllable identification tasks, in the presence of masking noise (Meister et al., 2007; d'Ausilio et al., 2009) or when processing acoustically ambiguous syllables (Möttonen and Watkins, 2009). Interestingly, compared to the present study using auditory adaptation, two recent studies also provided evidence that use-induced motor plasticity, by means of tongue and lip motor training, can bias perceptual performance in auditory speech recognition (Sato et al., 2011) and that sensory-motor brain regions are sensitive to changes in response bias (Venezia et al., 2012). In a similar vein, previous work has shown that motor learning/compensation induced by online auditory feedback manipulation in a speech production task changes both the speech motor output and the auditory speech representations, changes that together act to reduce the impact of the altered feedback (Shiller et al., 2009). Consistent with these findings, our results strongly suggest that the dorsal pathway mediate speech decision/categorization processes and are consistent with recent neurobiological models of speech perception (Callan et al., 2004; Wilson and Iacoboni, 2006; Hickok and Poeppel, 2007; Skipper et al., 2007; Poeppel et al., 2008; Rauschecker and Scott, 2009; Hickok et al., 2011; Rauschecker, 2011). According to these models, processing speech sounds is assumed to depend upon successive sensory-to-motor and motor-to-sensory projections, with implicit procedural knowledge of speech production providing motor-based predictions (by the use of sensorimotor transformations and internal models; see Wolpert et al., 1995; Kawato, 1999; Wolpert and Flanagan, 2001; Guenther,

2006; Hickok et al., 2011; Guenther and Vladusich, 2012; Perkell, 2012). For example, Skipper et al. (2007) proposed that phonetic hypotheses are first derived from acoustico-phonetic analyses in the auditory system. These hypotheses are then mapped onto speech motor commands, which, in turn, may partly constrain phonetic interpretation by predicting the acoustic consequences of a speech movement through an efference copy to the auditory system. From that view, our finding that TMS combined with prior auditory adaptation can change the listener's sensorimotor state and thereby change the listener's classification of the syllable is consistent with these theoretical models.

Finally, although the present findings suggest an active role of the auditory dorsal pathway in speech categorization/decision processes when resolving stimulus ambiguity, it is worth noting that the question of whether articulatory processes mediate speech perception under normal listening conditions remains vigorously debated (e.g., Hickok and Poeppel, 2007; Meister et al., 2007; Schwartz et al., 2008; d'Ausilio et al., 2009, 2012; Lotto et al., 2009; Sato et al., 2009; Scott et al., 2009; Sato et al., 2011; Tremblay and Small, 2011). Indeed, several research findings do not support a mandatory role for the motor system in speech perception with clear auditory inputs. Firstly, damage to motor speech areas in Broca's aphasic patients does not produce clear deficits in speech perception, with aphasic patients performing well on auditory word comprehension tasks (e.g., Hickok et al., 2011). Results from both electrocortical stimulation studies during neurosurgical operations or from transcranial magnetic stimulation (TMS) also challenge a possible functional role of the motor system in speech processing under normal listening conditions (for reviews, see Boatmann, 2004; Sato et al., 2009; Scott et al., 2009). Indeed, temporarily disrupting the activity of the opercular part of the left pIFG or the PMv has been shown to disrupt subjects' ability to perform sublexical phonological tasks that require strong segmentation processes and working memory demands (Boatmann, 2004; Nixon et al., 2004; Romero et al., 2006; Sato et al., 2009). However, no interference effects were observed in non-word syllable identification/discrimination tasks that could be performed without phonemic segmentation (Boatmann, 2004; Boatman and Miglioretti, 2005; Sato et al., 2009; d'Ausilio et al., 2012). Hence, while the present findings do not speak to the mediating role of the auditory dorsal stream in speech perception under normal listening conditions, they confirm a causal role for several regions of the dorsal stream in selective adaption to speech in a simple discrimination task.

4. Conclusion

The present state-dependent transcranial magnetic stimulation study aimed at determining whether speech recognition is mediated by the auditory dorsal pathway, by examining the causal contribution of the left ventral premotor cortex, supramarginal gyrus and posterior part of the superior temporal gyrus in an auditory syllable identification/categorization task. To this aim, auditory perceptual adaptation was first behaviorally induced, with the goal of tuning specific neural populations in the targeted brain areas. As expected, in a subsequent auditory syllable decision task on ambiguous syllables, a strong selective adaption was observed as compared to a control condition without prior auditory adaptation. Crucially, compared to the results observed without stimulation, single-pulse transcranial magnetic stimulation delivered at the onset of each target stimulus interacted with the initial state of each of the stimulated brain area by enhancing the adaptation effect. These results suggest that regions of the dorsal streams are part of a network that contributes to auditory speech adaptation and provide further evidence for sensorimotor interaction in speech, and for a mediating role of the auditory dorsal pathway in speech categorization.

5. Experimental procedures

5.1. *Participants*

Twelve healthy native Canadian French speakers participated in the study (9 females; mean age ± SD: 27 ± 2 years). All participants were right-handed (Oldfield, 1971), had normal or corrected-to-normal vision and reported no history of language or hearing disorders. Participants were screened for speech/language disorders, neurological, psychiatric, and other medical conditions, as well as contraindications to TMS (Wassermann, 1998). Written informed consent was obtained for all participants; participants were compensated for the time spent in the study. The study was approved by the Magnetic Resonance Research Committee (MRRC) and the Montreal Neurological Institute (MNI) Research Ethics Committee.

5.2. *Stimuli*

For the categorical decision tasks performed without stimulation, the stimuli consisted of eleven synthesized syllables varying along a /ba/–/da/ continuum. These stimuli were generated using a pitch-modulated sinewave synthesis method adapted from a previous study on categorical perception (Serniclaes et al., 2001). More specifically, this consisted of generating a sum of three sinewaves at the formant frequency F_1, F_2 and F_3, and to modulate this sum with a pitch-synchronous comb envelope with an arbitrary pattern (a negative exponential function with a time constant of 10 ms) as well as to control the fundamental frequency F_0. The endpoints were given appropriate values for the perception of a French /ba/ syllable, at one end, and for the perception of a French /da/ syllable at the other end. The difference in place of articulation between /ba/ and/da/ syllables was created by modifying the onset of the initial frequency transitions (FT_2 and FT_3) which corresponded to those of the second and third formants (F_2 and F_3) in natural speech. The FT_2 onset frequency varied from 1150 Hz at the /ba/ endpoint to 1450 Hz at the /da/ endpoint in ten equal steps of 30 Hz. The FT_3 onset frequency varied from 2250 Hz at the /ba/ endpoint to 2750 Hz at the /da/ endpoint in ten equal steps of 50 Hz. This yielded a total of eleven stimuli per continuum (c^{-5} to c^{+5}). The end frequencies of FT_2 and FT_3

transition were fixed at 1300 Hz and 2500 Hz respectively (corresponding to the onset frequencies for c_0). For every stimulus, the initial frequency of the first formant (F_1) was 200 Hz and its end frequency was 750 Hz. For each stimulus, the amplitudes of the F_1, F_2, and F_3 sinewaves varied linearly from respectively .1, .001, .0001 to .3, .15, and .2 to simulate the respective contribution of the formants in speech signal energy for such type of syllables. The negative voice onset time (VOT) duration was 55 ms, the duration of all frequency transitions was 30 ms, and the duration of the stable vocalic segment was 130 ms. During negative VOT and frequency transition period, the fundamental frequency F_0 was fixed to 100 Hz. During the stable vocalic segment, a quasi-linear descending F_0 contour was applied (from 106 Hz to 92 Hz).

In the two forced choice (2FC) syllable decision tasks, the three syllables (c^-, c^0, c^+) ranging in the categorical boundary between /ba/ and/da/ were individually determined from the initial categorical decision task (see below).

5.3. *Procedure*

The experiment was carried out in a quiet room. Participants sat in front of a computer monitor at a distance of approximately 50 cm. The acoustic stimuli were presented at a comfortable sound level through headphones. Presentation software (Neurobehavioral Systems, Albany, CA) was used to control the stimulus presentation and to record keyboard presses.

After MRI-to-head co-registration and resting motor threshold (RMT) determination (see below), participants underwent a categorical perception task (without TMS), three 2FC syllable decision tasks (using TMS and according to each stimulation site: pSTG, SMG, vPM) and a second categorical perception task (without TMS). The experimental design is illustrated in Fig. 1.

5.3.1. *Categorical decision tasks*

In order to individually determine the three syllables (c^-, c^0, c^+) ranging in the categorical boundary between /ba/ and/da/, participants first performed a categorical decision task, performed without stimulation, on the eleven stimuli of the ba/–/da/ continuum (c^{-5} to c^{+5}). Each trial started with a fixation cue (the '+' symbol) presented in the middle of the screen for 500 ms, followed by the presentation of a syllable, and ended with a blank screen for 1500 ms. Participants were instructed to produce a motor response by pressing with their right index or middle finger one of two keyboard keys corresponding to /ba/ or/da/ syllable. Every stimulus was presented 20 times in a pseudo-randomized sequence (the same syllable never occurring twice in succession) for a total of 220 trials. In order to compensate for the acoustic artefacts present during stimulation in the subsequent two forced-choice (2FC) syllable decision tasks (due to single pulse clicking sound, see below), a single-pulse TMS was applied at the onset of each stimulus with the TMS coil turned and moved 10 cm away from participant's head during the categorical perception task. The three stimuli (c^-, c^0, c^+) best fitting the categorical boundary between /ba/ and/da/ were determined for each participant. In order to test for a possible modulation of categorical boundary due to stimulation in the two forced choice (2FC) syllable decision tasks, a second categorical perception task was performed after the TMS experiment, using the same experimental procedure.

5.3.2. *2FC syllable decision tasks*

Three 2FC syllable decision tasks were performed, one for each stimulation site (pSTG, SMG, vPM). With the exception of the stimulation site, the experimental procedure was identical in the three experimental sessions. For each stimulation site, four conditions were contrasted and related to the auditory adaptation mode (ADAPTATION, CONTROL) and the stimulation mode (TMS, SHAM). The order of the three syllable decision tasks (stimulation site), the stimulation mode, the adaptation mode and the response key designation (/ba/ vs./da/) were counterbalanced across participants.

For each stimulation site, each condition was presented twice (BIN1, BIN2) for a total of eight experimental blocks (see Fig. 1). Each block began with either an adaptation or a control task (40 s) and was followed by a syllable decision task (60 s). In four of these blocks (ADAPTATION condition), participants listened to a sequence of 40 /ba/ (corresponding to the /ba/ endpoint of the auditory continuum) presented at a rate of 1 Hz before performing the syllable decision task. In the other four blocks (CONTROL condition), a 40 s resting period without any auditory stimulation preceded the syllable decision task. The syllable decision tasks were performed on c^-, c_0 and c^+ stimuli. Each trial started with a fixation cue (the '+' symbol) presented in the middle of the screen for 500 ms, which was followed by the presentation of a syllable. The trial ended with a blank screen presented for 1500 ms. Participants were instructed to produce a motor response as quickly and accurately as possible, by pressing on one of two keys corresponding to /ba/ or/da on a computer keyboard with either their right index or middle finger. Every stimulus (c^-, c_0 or c^+) was presented 10 times in a pseudo-randomized sequence for a total of 30 ambiguous stimuli per condition. In four blocks (TMS condition), a single-pulse TMS was applied at the onset of each stimulus. In order to provide identical acoustic artefacts (due to single pulse clicking sound in the TMS condition) but without stimulation, the TMS coil was turned and moved 10 cm away from participant's head in the four other blocks (SHAM condition).

In sum, for each stimulation site (pSTG, SMG, vPM), the TMS and SHAM conditions each occurred in four successive blocks, with the ADAPTATION and CONTROL conditions occurring once in both the first and last two blocks (BIN1, BIN2). The order of all these conditions was counterbalanced across participants. Each experimental session (pSTG, SMG, vPM) thus consisted of 240 trials with 30 trials per experimental conditions: stimulation mode (TMS, SHAM) × adaptation mode (ADAPTATION, CONTROL) × bin (BIN1, BIN2).

5.4. *Transcranial magnetic stimulation*

5.4.1. *MRI acquisition and co-registration*

A high-resolution T1-weighted structural volume was acquired for anatomical localization for each participant. Data were acquired on a 1.5T Siemens Sonata MR scanner at the Montreal Neurological Institute (matrix 256 × 256 mm2, 176 slices, 1 × 1 × 1 mm3, no gap, TE=9.2 ms, TR=22 ms, flip

angle 30°). Once obtained, the anatomical MRI was incorporated into BrainSight 2 (Rogue Research, Montreal, Canada) to guide coil placement. Stimulation sites were identified on individual brain reconstructions on the basis of macroanatomical landmarks. pSTG was defined as the ventral region over the caudal end of the posterior branch of the sylvian fissure. SMG was defined as its dorsalmost portion, immediately ventral to the intraparietal sulcus. vPM was defined as the portion of the precentral gyrus posterior to the point where the inferior frontal sulcus intersects with the precentral sulcus (see Fig. 1D). The mean coordinates for all participants in MNI space were x=−56 (±.9), y=−37 (±2.3), z=12 (±1.2) for pSTG, x=−47 (±1), y=−49 (±2.2), z=38 (±1.6) for SMG and x=−47 (±.8), y=5 (±2), z=33 (±1) for vPM.

Prior to the experiment, an MRI-to-head co-registration was performed for each participant. The position of four anatomical landmarks (tip of the nose, bridge of the nose, superior-lateral edge of the tragus of left and right ears), previously identified on participant's MRI, was assessed using an infrared tracking system (Polaris, Northern Digital, Waterloo, Canada). Upon successful co-registration, infrared tracking was used to monitor the position of the coil with respect to the participant's brain and the stimulation site.

5.4.2. *Resting motor threshold (RMT)*

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). For the determination of each participant's RMT, the TMS coil was first placed over the participant's left hand primary motor cortex with the coil held tangentially to the skull, and the handle pointing posteriorly and inferiorly. The location of the stimulation was then adjusted to locate the maximally excitable hand area. RMT was established as the lowest stimulation intensity applied over the hand primary motor cortex capable of evoking a contraction in the relaxed right hand muscles on at least 5 out of 10 consecutive stimulations.

5.4.3. *TMS experiment*

For each syllable decision task, the TMS coil was first positioned on the stimulation site (pSTG, SMG, vPM) using frameless stereotaxy and individual MRI-to-head co-registration. The coil was held tangentially to the skull and secured by a positioning arm. Coil orientation was antero-posterior with the handle pointing backward for vPM, perpendicular to the midline with the handle pointing outward for pSTG and SMG. During the tasks, its position was monitored online and adjusted following any head movements resulting in displacement of more than 2 mm in any direction to the stimulation site. In all conditions, a single TMS pulse was applied at the onset of each auditory stimulus, with stimulation intensity at 110% of individual RMT (mean value of stimulation of 70%, ±2%). In the SHAM condition, the TMS coil was turned and moved 10 cm away from participant's head in order to provide identical acoustic artifacts (due to single pulse clicking sound) but without stimulation.

5.5. *Data analysis*

For all the following analyses, the significance level was set at $p=.05$ and Greenhouse–Geisser corrected (for violation of the sphericity assumption) when appropriate. When required, posthoc analyses were conducted with protected LSD Fischer tests.

5.5.1. *Categorical decision task*

The proportion of /ba/ responses observed in the categorical decision tasks was computed for each participant as a function of the eleven stimuli of the /ba/–/da/ continuum (c^{-5} to c^{+5}) in each experimental session (before and after the TMS experiment). A two-way repeated-measure analysis of variance (ANOVA) was performed on the proportion of /ba/ responses with the stimulus and the experimental session as within-subjects variables.

5.5.2. *2FC syllable decision task*

Missing responses and trials in which reaction-times (RTs), calculated from the stimulus onset, were faster than 300 ms (anticipations) or slower than 1500 ms (missing responses) were considered as errors and discarded without replacement. Two participants were removed from the analysis because their error rates exceeded 10%. For the other ten participants, the mean percentage of errors was 2% ±1%. The proportion of /ba/ responses and the median RTs (Whelan 2008) observed in the TMS experiment were then computed for each participant, each stimulation site (pSTG, SMG, vPM), each stimulation mode (TMS, SHAM), each adaptation mode (ADAPTATION, CONTROL) and each bin (BIN1, BIN2). To test the effect of the different experimental conditions on selective adaptation, four-way repeated-measure ANOVAs were performed on these two dependent variables with stimulation site, stimulation mode, adaptation mode and bin as within-subjects variables.

Acknowledgments

The authors declare no conflict of interest. This study was supported by research grants from CNRS (Centre National de la Recherche Scientifique) and ANR (Agence Nationale de la Recherche, ANR SPIM 'Imitation in Speech') to MS and from PPF (Plan Pluri-formations "Interactions Multimodales") to KG. We thank Lionel Granjon and Willy Serniclaes for their help in designing the stimuli. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies.

REFERENCES

Alho, J., Sato, M., Sams, M., Schwartz, J.-L., Tiitinen, H., Jaaskelainen, I.P., 2012. Enhanced early-latency electromagnetic activity in the left premotor cortex is associated with successful phonetic categorization. NeuroImage 60 (4), 1937–1946.

Boatmann, 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. J. Neurosci. 25 (23), 5475–5480.
Callan, D., Callan, A., Gamez, M., Sato, M.A., Kawato, M., 2010. Premotor cortex mediates perceptual performance. NeuroImage 51 (2), 844–858.
Callan, D.E., Jones, J.A., Callan, A.M., 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.
Callan, D.E., Jones, J.A., Munhall, K.G., Callan, A.M., Kroos, C., Vatikiotis-Bateson, E., 2003. Neural processes underlying perceptual enhancement by visual speech gestures. Neuroreport 14, 2213–2217.
Calvert, G.A, Campbell, R., 2003. Reading speech from still and moving faces: the neural substrates of visible speech. J. Cognitive Neurosci. 15 (1), 57–70.
Cattaneo, L., 2010. Tuning of ventral premotor cortex neurons to distinct observed grasp types: a TMS-priming study. Exp. Brain Res. 207 (3–4), 165–172.
Cattaneo, L., Barchiesi, G., Tabarelli, D., Arfeller, C., Sato, M., Glenberg, A.M., 2011. One's motor performance predictably modulates the understanding of others' actions through adaptation of premotor visuo-motor neurons. Soc. Cognitive Affective Neurosci. 6 (3), 301–310.
Cattaneo, L., Sandrini, M., Schwarzbach, J., 2010. State-dependent TMS reveals a hierarchical representation of observed acts in the temporal, parietal, and premotor cortices. Cereb. Cortex 20 (9), 2252–2258.
Cattaneo, Z., Silvanto, J., 2008. Investigating visualmotion perception using the transcranial magnetic stimulation–adaptation paradigm. Neuroreport 19, 1423–1427.
Cattaneo, Z., Devlin, J.T., Salivini, F., Vecchi, T., Silvanto, J., 2010. The causal role of category-specific neuronal representations in the left ventral premotor cortex, PMv in semantic processing. NeuroImage 49 (3), 2728–2734.
Chong, T.T, Cunnington, R., Williams, M.A., Kanwisher, N., Mattingley, J.B., 2008. fMRI adaptation reveals mirror neurons in human inferior parietal cortex. Curr. Biol. 18 (20), 1576–1580.
Cooper, W., 1979. Speech Perception and Production: Studies in Selective Adaptation. Ablex, Norwood, NJ.
Cooper, W.E., Lauritsen, M.S., 1974. Feature processing in the perception and production of speech. Nature 252, 121–123.
d'Ausilio, A., Craighero, L., Fadiga, L., 2012. The contribution of the frontal lobe to the perception of speech and language. J. Neurolinguist. 25 (4), 328–335.
d'Ausilio, A., Bufalari, I., Salmas, P., Fadiga, L., 2012. The role of the motor system in discriminating degraded speech sounds. Cortex 48 (7), 882–887.
d'Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., Fadiga, L., 2009. The motor somatotopy of speech perception. Curr. Biol. 19 (5), 381–385.
Dick, A.S., Tremblay, P., 2012. Beyond the arcuate fasciculus: consensus and controversy in the connectional anatomy of language. Brain 135 (12), 3529–3550.
Dinstein, I., Hasson, U., Rubin, N., Heeger, D.J., 2007. Brain areas selective for both observed and executed movements. J. Neurophysiol. 98 (3), 1415–1427.
Eimas, P.D., Corbit, J.D., 1973. Selective adaptation of linguistic feature detectors. Cognitive Psychol. 4, 99–109.
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.
Galantucci, B., Fowler, C.A., Turvey, M.T., 2006. The motor theory of speech perception reviewed. Psychon. Bull. Rev. 13 (3), 361–377.
Glenberg, A.M., Lopez-Mobilia, G., McBeath, M., Toma, M., Sato, M., Cattaneo, L., 2010. Knowing beans: Human mirror mechanisms revealed through motor adaptation. Front. Human Neurosci. 4, 206.
Gough, P.M., Nobre, A.C., Devlin, J.T., 2005. Dissociating linguistic processes in the left inferior frontal cortex with transcranial magnetic stimulation. J. Neurosci 25, 8010–8016.
Grabski, K., Schwartz, J.-L., Lamalle, L., Vilain, C., Vallée, N., Baciu, M., Le Bas, J.-F,, Sato, M., 2013. Shared and distinct neural correlates of vowel perception and production. J. Neurolinguist (3), 384–408.
Grill-Spector, K., Henson, R., Martin, A., 2006. Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn. Sci. 10 (1), 14–23.
Grill-Spector, K., Malach, R, 2001. fMR-adaptation: a tool for studying the functional properties of human cortical neurons. Acta Psychol. Amst. 107 (1–3), 293–321.
Guenther, F.H., Vladusich, T., 2012. A neural theory of speech acquisition and production. J. Neurolinguist. 25 (4), 408–422.
Guenther, F.H., 2006. Cortical interactions underlying the production of speech sounds. J. Commun. Disord. 39, 350–365.
Helmholtz, H., 1867. Handbuch der physiologischen Optik. L. Voss, Leipzig.
Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nat. Rev. Neurosci. 8, 393–402.
Hickok, G., Houde, J., Rong, F., 2011. Sensorimotor integration in speech processing: computational basis and neural organization. Neuron 69 (3), 407–422.
Hickok, G., Costanzo, M., Capasso, R., Miceli, G., 2011. The role of Broca's area in speech perception: evidence from aphasia revisited. Brain Lang. 119 (3), 214–220.
Kawato, M., 1999. Internal models for motor control and trajectory planning. Curr. Opin. Neurobiol. 9 (6), 718–727.
Kilner, J.M., Neal, A., Weiskopf, N., Friston, K.J., Frith, C.D., 2009. Evidence of mirror neurons in human inferior frontal gyrus. J. Neurosci. 29 (32), 10153–10159.
Liberman, A.M, Whalen, D.H., 2000. On the relation of speech to language. Trends Cognitive Sci. 3 (7), 254–264.
Liberman, A.M., Mattingly, I.G., 1985. The motor theory of speech perception revised. Cognition 21, 1–36.
Liberman, A.M., Cooper, F.S., Shankweiler, D.P., Studdert-Kennedy, M., 1967. Perception of the speech code. Psychol. Rev. 74, 431–461.
Lingnau, A., Gesierich, B., Caramazza, A., 2009. Asymmetric fMRI adaptation reveals no evidence for mirror neurons in humans. Proc. Natl. Acad. Sci. USA 106, 9925–9930.
Lotto, A., Hickok, G., Holt, L., 2009. Reflections on mirror neurons and speech perception. Trends Cogn Sci. 13, 110–114.
Meister, I.G., Wilson, S.M., Deblieck, C., Wu, A.D., Iacoboni, M., 2007. The essential role of premotor cortex in speech perception. Curr. Biol. 17 (19), 1692–1696.
Möttönen, R., Watkins, K.E., 2009. Motor representations of articulators contribute to categorical perception of speech sounds. The Journal of Neuroscience (31), 29–9825.
Nikouline, V., Ruohonen, J., Ilmoniemi, R.J., 1999. The role of the coil click in TMS assessed with simultaneous EEG. Clin. Neurophysiol. 110, 1325–1328.
Nishitani, N., Hari, R., 2002. Viewing lip forms: cortical dynamics. Neuron 36, 1211–1220.
Nixon, P., Lazarova, J., Hodinott-Hill, I., Gough, P., Passingham, R., 2004. The inferior frontal gyrus and phonological processing: an investigation using rTMS. J Cogn Neurosci 16 (2), 289–300.
Ojanen, V., Möttönen, R., Pekkola, J., Jääskeläinen, I.P., Joensuu, R., Autti, T., Sams, M., 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.

Paulesu, E., Perani, D., Blasi, V., Silani, G., Borghese, A.A., De Giovanni, U., Sensolo, S., Fazio, F., 2003. A functional-anatomical modelfor lipreading. J. Neurophysiol. 90, 2005–2013.
Pekkola, J., Laasonen, M., Ojanen, V., Autti, T., Jaaskelainen, L.P., Kujala, T., Sams, M., 2006. Perception of matching and conflicting audiovisual speech in dyslexic and fluent readers: an fMRI study at 3T. NeuroImage 29 (3), 797–807.
Perkell, J.S., 2012. Movement goals and feedback and feedforward control mechanisms in speech production. J. Neurolinguist. 25 (4), 382–407.
Poeppel, D., Idsardi, W.J., van Wassenhove, V., 2008. Speech perception at the interface of neurobiology and linguistics. Philos. Trans. R. Soc. London B Biol. Sci. 363 (1493), 1071–1086.
Pulvermuller, F., Huss, M., Kherif, F., Martin, Moscoso del Prado, Hauk, F., Shtyrov, Y., O., 2006. Motor cortex maps articulatory features of speech sounds. Proc. Natl. Acad. Sci. USA 103 (20), 7865–7870.
Rauschecker, J.P., 2011. An expanded role for the dorsal auditory pathway in sensorimotor control and integration. Hear. Res. 271, 16–25.
Rauschecker, J.P., Scott, S.K., 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nat. Neurosci. 12 (6), 718–724.
Rizzolatti, G, Craighero, L., 2004. The mirror–neuron system. Annu. Rev. Neurosci. 27, 169–192.
Rizzolatti, G., Fogassi, L., Gallese, V., 2001. Neurophysiological mechanisms underlying the understanding and imitation of action. Nat. Rev. Neurosci. 2, 661–670.
Romero, L., Walsh, V., Papagno, C., 2006. The neural correlates of phonological short-term memory: a repetitive transcranial magnetic stimulation study. J. Cognitive Neurosci. 18 (7), 1147–1155.
Roy, A.C., Craighero, L., Fabbri-Destro, M., Fadiga, L., 2008. Phonological and lexical motor facilitation during speech listening: a transcranial magnetic stimulation study. J. Physiol. Paris 102 (1–3), 101–105.
Samuel, A.G., 1986. Red herring detectors and speech perception: In defense of selective adaptation. Cognitive Psychology 18, 452–499.
Sato, M., Buccino, G., Gentilucci, M., Cattaneo, L., 2010. On the tip of the tongue: modulation of the primary motor cortex during audiovisual speech perception. Speech Commun. 52 (6), 533–541.
Sato, M., Grabski, K., Glenberg, A., Brisebois, A., Basirat, A., Ménard, L., Cattaneo, L., 2011. Articulatory bias in speech categorization: evidence from use-induced motor plasticity. Cortex 47 (8), 1001–1003.
Sato, M., Tremblay, P., Gracco, V., 2009. A mediating role of the premotor cortex in phoneme segmentation. Brain Lang. 111 (1), 1–7.
Schwartz, J.-L., Abry, C., Boë, L.-J., Cathiard, M.A., 2002. Phonology in a theory of perception-for-action-control. In: Durand, J., Lacks, B. (Eds.), Phonology: From Phonetics to Cognition. Oxford University Press, Oxford, pp. 240–280.
Schwartz, J.-L., Ménard, L., Basirat, A., Sato, M., 2012. The perception for action control theory, PACT: a perceptuo-motor theory of speech perception. J. Neurolinguist. 25 (4), 336–354.
Schwartz, J.-L., Sato, M., Fadiga, L., 2008. The common language of speech perception and action: a neurocognitive perspective. Rev. Fr. Linguist. Appl. 13 (2), 9–22.
Scott, S.K., 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 (4), 295–302.
Serniclaes, W., Sprenger-Chraolles, L., Carré, R., Démonet, J-.F., 2001. Perceptual discrimination of speech sounds in developmental dyslexia. J. Speech Lang. Hear. Res. 44 (2), 384–399.
Shiller, D., Sato, M., Gracco, V., Baum, S., 2009. Perceptual recalibration of speech sounds following speech motor learning. J. Acoust. Soc. Am. 125 (2), 1103–1113.
Silvanto, J., Muggleton, N., Walsh, V., 2008. State-dependency in brain stimulation studies of perception and cognition. Sci. 12, 447–454.
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. Cereb. Cortex 17 (10), 2387–2399.
Stevens, K.N., Halle, M., 1967. Remarks on analysis by synthesis and distinctive features. In: Walthen-Dunn, W. (Ed.), Models for the Perception of Speech and Visual Form. MIT Press, Cambridge, MA, pp. 88–102.
Sundara, M., Namasivayam, A.K., Chen, R., 2001. Observation-execution matching system for speech: a magnetic stimulation study. Neuroreport 12 (7), 1341–1344.
Tremblay, P., Small, S.L., 2011. On the context-dependent nature of the contribution of the ventral premotor cortex to speech perception. NeuroImage 57, 1561–1571.
Venezia, J.H., Saberi, K., Chubb, C., Hickok, G., 2012. Response bias modulates the speech motor system during syllable discrimination. Front. Psychol. 3, 157.
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., Electroencephalogr Clin Neurophysiol 108 (1), 1–16.
Watkins, K.E., Paus, T., 2004. Modulation of motor excitability during speech perception: the role of Broca's area. J. Cognitive Neurosci. 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.
Whelan, 2008. Effective analysis of reaction time data. Psychol. Rec. 58, 475–482.
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. Nat. Neurosci. 7, 701–702.
Wolpert, D.M., Flanagan, J.R., 2001. Motor prediction. Curr. Biol. 18, R729–R732.
Wolpert, D.M., Ghahramani, Z., Jordan, M.I., 1995. An internal model for sensorimotor integration. Science 269, 1880–1882.