

ORIGINAL ARTICLE

Causal Influence of Articulatory Motor Cortex on Comprehending Single Spoken Words: TMS Evidence

Malte R. Schomers^{1,2}, Evgeniya Kirilina³, Anne Weigand^{3,4,5}, Malek Bajbouj^{3,4}, and Friedemann Pulvermüller^{1,2}

¹Brain Language Laboratory, Department of Philosophy and Humanities, Freie Universität Berlin, 14195 Berlin, Germany, ²Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, 10099 Berlin, Germany, ³Dahlem Institute for Neuroimaging of Emotion, Freie Universität Berlin, 14195 Berlin, Germany, ⁴Department of Psychiatry, Charité Universitätsmedizin Berlin, 14050 Berlin, Germany, and ⁵Berenson-Allen Center for Noninvasive Brain Stimulation, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA 02215, USA

Address correspondence to Friedemann Pulvermüller, Brain Language Laboratory, Department of Philosophy and Humanities, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany. Email: friedemann.pulvermuller@fu-berlin.de

Abstract

Classic wisdom had been that motor and premotor cortex contribute to motor execution but not to higher cognition and language comprehension. In contrast, mounting evidence from neuroimaging, patient research, and transcranial magnetic stimulation (TMS) suggest sensorimotor interaction and, specifically, that the articulatory motor cortex is important for classifying meaningless speech sounds into phonemic categories. However, whether these findings speak to the comprehension issue is unclear, because language comprehension does not require explicit phonemic classification and previous results may therefore relate to factors alien to semantic understanding. We here used the standard psycholinguistic test of spoken word comprehension, the word-to-picture-matching task, and concordant TMS to articulatory motor cortex. TMS pulses were applied to primary motor cortex controlling either the lips or the tongue as subjects heard critical word stimuli starting with bilabial lip-related or alveolar tongue-related stop consonants (e.g., “pool” or “tool”). A significant cross-over interaction showed that articulatory motor cortex stimulation delayed comprehension responses for phonologically incongruent words relative to congruous ones (i.e., lip area TMS delayed “tool” relative to “pool” responses). As local TMS to articulatory motor areas differentially delays the comprehension of phonologically incongruous spoken words, we conclude that motor systems can take a causal role in semantic comprehension and, hence, higher cognition.

Key words: action-perception theory, language comprehension, motor system, speech processing, transcranial magnetic stimulation

Introduction

One of the most fundamental debates in current neuroscience addresses the role of the frontal lobe in perception and understanding. The mirror neuron literature has shown action-perception

linkage and premotor activation in perceptual and comprehension processes (Rizzolatti et al. 1996, 2014), but to which degree this frontal activation is necessary for understanding remains controversial. The left frontal cortex’ contribution to symbol understanding has long been in the focus of brain research on

language. In neurolinguistics, the classic view had been that the left inferior frontal cortex, including Broca's area (Brodmann areas [BA] 44, 45), and adjacent articulatory motor areas, including inferior motor and premotor cortex (BA 4, 6), are involved in speech production only, whereas for speech perception and language comprehension, auditory and adjacent Wernicke's area in superior-temporal cortex (including BA 41, 42, 22) are necessary and sufficient (Wernicke 1874; Lichtheim 1885; Geschwind 1970). In sharp contrast with this position, action-perception theory of language postulates that the neuronal machineries for speech production and understanding are intrinsically linked with each other, implying a double function of both superior-temporal and, critically, inferior-frontal cortex in both production and comprehension (Pulvermüller and Fadiga 2010). A crucial prediction of the latter perspective, which goes against the classic approach, is that left inferior frontal and even articulatory motor areas are causally involved in language comprehension, including the understanding of single words.

Several arguments seem to support a role in comprehension of left inferior frontal cortex and adjacent articulatory motor systems. Liberman and his colleagues pointed out in the context of their "motor theory of speech perception" that there is great context-dependent variability of acoustic-phonetic features of speech sounds that fall in the same phonemic category and that a link between these disparate auditory schemas might be possible based on the similar articulatory gestures performed to elicit the sounds (Liberman and Mattingly 1985; Galantucci et al. 2006). In view of the neurobiological mechanisms of speech perception, it was argued that the correlated articulatory motor and auditory information available to the human brain during babbling and word production fosters the development of action-perception links between frontal articulatory and temporal auditory areas. Formation of such frontotemporal links for phonemes is implied by the fundamental principle of Hebbian correlation learning and the neuroanatomical connectivity profile between frontal and temporal cortex (Rilling et al. 2008), thus yielding functional frontotemporal connections between neurons controlling articulatory motor movements performed to produce speech sounds and neurons involved in acoustic perception of the corresponding speech sounds (Braitenberg and Pulvermüller 1992; Pulvermüller et al. 2006; Pulvermüller and Fadiga 2010). In this view, neuronal circuits distributed across superior-temporal and inferior-frontal areas that reach into articulatory motor cortex carry both language production and comprehension. Further arguments against a functional separation of receptive and productive language areas come from studies in patients with post-stroke aphasia. Broca's aphasics, in addition to their speech production deficits, appear to be impaired in speech perception tasks, which typically involve the explicit discrimination between, or identification of, phonemes or syllables (Basso et al. 1977; Blumstein et al. 1977, 1994; Utman et al. 2001). However, attempts at replicating these findings were not always successful and, crucially, as critical lesions are sometimes extensive, lesion studies do not provide unambiguous evidence for a general involvement of left inferior frontal cortex in speech perception. For example, only three Broca's aphasics in a group study documenting a perception deficit indeed had lesions restricted to frontoparietal cortex (Utman et al. 2001); whereas one study (Rogalsky et al. 2011) failed to detect profound speech perception deficits in two aphasics with selective inferior frontal lesions, another study (Caplan et al. 1995) documented them in a different set of two such patients.

In this context, it is important to emphasize the distinction between the *perception* of phonemes and the *comprehension* of

meaningful language. Although all linguistic signs serve the role of communicating meaning, phonemes, the smallest units that *distinguish between* meaningful signs, do themselves not *carry* meaning. If they are presented outside the context of words, they lack their normal function as meaning-discriminating units. Importantly, speech perception performance of aphasia patients dissociates from their ability to understand the meaning of single words (Miceli et al. 1980). Therefore, a deficit in a speech perception task does not imply a comprehension failure. Although some authors have argued that word comprehension is still relatively intact in some patients with aphasia, because semantic evidence can in part be used to reconstruct missing phonetic information (Basso et al. 1977; Blumstein et al. 1994), it is possible that word comprehension experiments reveal additional information about left inferior frontal cortex function. Spoken word comprehension deficits in patients with Broca's aphasia have been documented when stimuli were degraded and embedded in noise but not when they were spoken clearly (Moineau et al. 2005). However, a significant delay in comprehending clearly spoken single words was reported in this and related studies (Utman et al. 2001; Yee et al. 2008). Still, these delays may be attributable to general cognitive impairments or strategic aspects of language processing. Clear evidence from small well-documented lesions in left inferior frontal cortex documents comprehension deficits for specific semantic types of words, especially action words (Neininger and Pulvermüller 2001, 2003; Kemmerer et al. 2012), and a similar category-specific deficit is present in patients with degenerative brain diseases primarily affecting cortical regions in the frontal lobe, including Motor Neuron Disease and Parkinson's Disease (Bak et al. 2001; Boulenger et al. 2008). In summary, patient studies could so far not finally clarify whether the left inferior frontal cortex or the articulatory motor system serve a general causal role in speech comprehension. Any firm conclusions are hampered by uncertainties about precise lesion sites, known differences between speech perception and comprehension, uncertainties about the interpretation of delays in speech comprehension and the category-specific nature of some well-documented comprehension deficits.

In addition to patient studies, neuroimaging studies employing a wide variety of methods (univariate and multivariate fMRI, EEG/MEG, and connectivity analysis) have found activation of left inferior frontal and articulatory motor cortices in speech perception (Pulvermüller et al. 2003, 2006; Wilson et al. 2004; Osnes et al. 2011; Lee et al. 2012; Chevillet et al. 2013; Liebenthal et al. 2013; Alho et al. 2014; Du et al. 2014). One study by Pulvermüller et al. (2006) even revealed that information about the specific place of articulation of a passively perceived speech sound is manifest in focal activation of the articulatory representations in the motor system. In addition, transcranial magnetic stimulation (TMS)-evoked activation of articulatory muscles (lip or tongue) was shown to be increased while listening to meaningful speech (Fadiga et al. 2002; Watkins et al. 2003) and even when viewing hand gestures associated with spoken words (Komeilipoor et al. 2014). However, although these results show that speech perception and comprehension elicit left inferior frontal and articulatory motor cortex activation, the possibility exists that such activation is consequent to but not critical to perception and comprehension processes (Mahon and Caramazza 2008).

TMS can also be used to reveal a possible causal role of motor areas in language processing, thereby overcoming some of the limitations mentioned earlier. TMS induces functional changes in cortical loci, which can be localized with an accuracy of 5–10 mm (Walsh and Cowey 2000). This neuropsychological research strategy has been applied using tasks requiring explicit

classification of speech sounds embedded in meaningless syllables (Meister et al. 2007; D'Ausilio et al. 2009; Möttönen and Watkins 2009; for reviews, see Möttönen and Watkins 2012; Murakami et al. 2013). The results showed a causal effect on the perceptual classification of noise-embedded speech sounds.

In this context, it is important to highlight once again that, in previous TMS studies, phonemes were presented in isolation; thus, the tasks did not entail the comprehension of the meaning of speech and, therefore, it remains unclear whether the influence of articulatory motor areas extends to normal speech comprehension and semantic understanding.

In summary, on the basis of existing data from patient, neuroimaging, and TMS studies, the role of inferior frontal and motor cortex in language comprehension remains controversial. Some authors deny any role of these areas in comprehension completely or acknowledge an influence only in artificial tasks and/or degraded listening conditions (Hickok 2009, 2014; Rogalsky et al. 2011). Others implicate only anterior areas, in particular anterior parts of Broca's area in semantic processing, whereas more posterior parts of frontal cortex (posterior Broca's area and premotor/motor cortex) are considered at most to play a role in phonological processing (Bookheimer 2002; Gough et al. 2005). However, the crucial hypothesis about the role of left inferior frontal, and most importantly, left articulatory motor cortex in meaningful spoken word comprehension, still awaits systematic experimental testing.

To clarify the left articulatory motor cortex's role in single word comprehension, we here used the standard psycholinguistic test of spoken language comprehension, the word-to-picture-matching task. Naturally spoken words, which were minimal pairs differing only in their initial phoneme (e.g., "pool" and "tool"), were used as critical word stimuli to ascertain that, in the context of the experiment, the initial phonemes served their normal role as meaning-discriminating units. All critical phonemes were either bilabial (lip-related) or alveolar (tongue-related) stop consonants. Immediately before onset of the spoken word, TMS was applied to left articulatory motor cortex, either to the lip or tongue representation. For word-to-picture matching, two images appeared immediately after the word and subjects had to indicate as fast as possible, which of them corresponded to the meaning of the word. Classic brain-language models predict no influence of motor cortex TMS on single word comprehension, whereas action-perception theory of language suggests a modulation of language comprehension performance specific to phonological word type and specific to TMS locus.

Materials and Methods

Participants

Thirteen monolingual native speakers of German (6 females) with a mean age of 22 years (range: 18–28 years) participated in the experiment for financial compensation or course credit. They had no history of neurological or psychiatric illness, normal or corrected-to-normal vision, and normal hearing (as assessed by a questionnaire). All participants were right-handed (Oldfield 1971) (laterality quotient $M = 88.8$, $SD = 19.4$). Participants provided written informed consent prior to participating in the study, and procedures were approved by the Ethics Committee of the Charité University Hospital, Berlin, Germany.

Stimuli

Stimuli were a total of 140 German words. Forty-four of those were critical experimental stimuli, and 96 were filler stimuli.

The critical word stimuli were 22 minimal pairs differing only in their initial phoneme (see Supplementary Table 1). In each pair, one word started with a bilabial (and therefore "lip-related") stop consonant ([b] or [p]), and the other with an alveolar ("tongue-related") stop consonant ([d] or [t]). We refer to these words as "lip words" and "tongue words", respectively. Nineteen of the 22 word pairs were also matched in the feature (voicing), yielding similar voice onset times. Tongue and lip words were matched for the following psycholinguistic variables (mean \pm SD) obtained from the dlexDB database (Heister et al. 2011): normalized word type frequency (lip words: $M = 18 \pm 28.9$; tongue words: $M = 81 \pm 192$), normalized average character bigram corpus frequency (lip words: $M = 48\,942 \pm 26\,130$; tongue words: $M = 57\,234 \pm 26\,907$) and normalized average character trigram corpus frequency (lip words: $M = 39\,609 \pm 26\,864$; tongue words: $M = 49\,216 \pm 27\,421$). All comparisons between lip words and tongue words were non-significant ($P > 0.05$).

We took efforts to distract subjects from the phonological nature of the present experiment and its focus on tongue- vs. lip-related phonemes. To this end, 96 irrelevant "filler" stimuli were added to the 44 experimental ones, including 24 semantically, but not phonologically related word pairs (e.g., "apple" vs. "cherry"), 12 minimal pairs differing in their initial phonemes, which were not bilabial vs. alveolar stops (e.g., "key" vs. "fee"), and 12 word pairs differing in the final phonemes (e.g., "gun" vs. "gum"). The full list of filler stimuli is provided in Supplementary Table 2.

All stimulus words were naturally spoken by a female native speaker of German during a single session in a soundproof chamber. They were recorded using an ATR1200 microphone (Audio-Technica Corporation), digitized (44.1 kHz sampling rate), and stored on disk. After recording, noise was removed and loudness of the each recording normalized to -20 dB SPL using "Audacity" (<http://audacity.sourceforge.net>). Average duration of stimuli was 432 ms ($SD \pm 50$ ms) for lip words and 445 ms ($SD \pm 41$ ms) for tongue words; a paired t-test showed no significant difference in duration.

For each stimulus word, a corresponding picture was chosen. The corresponding picture showed an object the word is typically used to speak about. Two of the experimental words were too abstract to find an obvious pictorial correlate of their meaning, and a circle was used as abstractness symbol so that the decision between pictorial alternatives had to be based on the distinction between one semantic match and one "abstractness" indicator. For the remaining 42 critical stimulus words, semantically related pictures could be produced easily.

TMS Methods

Structural MRI images were obtained for all participants (3T, Tim Trio Siemens, T1-weighted images, isotropic resolution $1 \times 1 \times 1$ mm³) and used for frameless stereotactic neuronavigation (eXimia Navigated Brain Stimulation, Nexstim). TMS pulses were generated by a focal biphasic figure-of-eight coil (eXimia 201383P). Coil position was maintained at roughly 90° to the central sulcus in the direction of the precentral gyrus. Using single TMS pulses, the lip representation in the left motor cortex was localized in each subject by measuring electromyographic activity in two surface electrodes attached to the right orbicularis oris muscle (as described in Möttönen and Watkins 2009), or, if this was not possible, the first dorsal interosseus (FDI) muscle was localized according to standard procedures (Rossini 1994). Target articulator loci not directly localized in an individual were calculated using the method described in a previous TMS study by D'Ausilio et al. (2009): Montreal Neurological Institute (MNI)

coordinates for the left lip and tongue representation from an earlier fMRI study (Pulvermüller et al. 2006) were converted to individual subjects' head space. In subjects for whom only the individual FDI location was available, the target coordinates were adjusted according to the difference between the actual FDI location and standard FDI coordinates (Niyazov et al. 2005) ($x, y, z = -37, -25, 58$). In those subjects where the lip representation could be localized, only the tongue coordinate was adjusted according to the difference between actual and projected lip location. The average MNI coordinates of the actual stimulation sites were as follows: lip MNI $x, y, z = -55.4, -9.2, 43.9$ and tongue MNI $x, y, z = -59.4, -7.4, 22.8$ (see Fig. 1). Thus, the distance between average actual stimulation sites and standard MNI peaks found in the fMRI study by Pulvermüller et al. (2006) and used by D'Ausilio et al. (2009) was 2.5 mm for the lip site (MNI peak: $x, y, z = -56, -8, 46$) and 3.5 mm for the tongue site (MNI peak: $x, y, z = -60, -10, 25$). The average distance between the two stimulation sites was 22.3 mm.

In those subjects in whom lip motor threshold could be determined ($n = 3$), stimulation intensity was 90% of the lip motor threshold. In subjects for whom only FDI motor threshold was available ($n = 10$), on average 100% of the FDI threshold (Rossini 1994) was used because higher intensities are normally required to magnetically stimulate the articulators. Average intensity was 34.6% of total stimulator output. TMS pulses (10 Hz, i.e., separated by 100 ms) were delivered before the onset of the spoken target word. Previous studies (Moliadze et al. 2003; Mottaghy et al. 2006) showed that facilitatory TMS effects emerge when the delay between TMS and critical stimulus onset is in the range from >100 to 500 ms. We chose a delay of 200 ms between the last pulse and spoken word onset to minimize acoustic interference of the clicking sounds accompanying TMS pulses with the spoken stimuli. We initially used 3 pulses to potentially further increase TMS efficacy (Kammer and Baumann 2010); however, after running 3 subjects, we noticed that this caused a risk of coil overheating and thus used 2 pulses for the remaining subjects.

Experimental Procedures

Each subject attended two TMS sessions, separated by at least 2 weeks. There were two blocks per session. In each block, 70 trials were administered, each with 22 critical trials and 48 filler trials. Trial order was pseudo-randomized with no more than three stimuli of one type allowed in direct succession. A trial consisted of one spoken word immediately followed by a picture pair. Subjects had to indicate as fast and as accurately as possible which of the pictures matched the spoken words. For each word pair, the two complementary spoken targets were presented in separate blocks of each session. Alternative pseudorandomization orders were produced using the software Mix (van Casteren and Davis 2006) so that all stimuli were identical in the two sessions but in a different order. In each block, TMS was delivered to one articulator locus, the order being counterbalanced over subjects. Stimulus word delivery was through in-ear headphones (Koss Corporation), which also provided attenuation of noise created by the TMS machine and stimulation.

To reduce the degree of redundancy immanent to the speech signal and avoid ceiling effects (see Discussion for further explanation), we individually adjusted the sound pressure level to a range where subjects were still able to repeat most of a set of test words correctly ($M = 69\%$, $SEM = 3.7\%$) in a pilot screening. This screening was followed by a training session consisting of 11 trials of word-to-picture matching but with different items than those used in the experiment. During the training session,

TMS pulses were also delivered so that subjects could familiarize themselves with the task, in particular with carrying it out under concurrent TMS stimulation.

Each trial started with a fixation cross in the middle of the screen displayed for an interval randomly varying between 1 to 2 s, after which the target sound was played via the headphones. The TMS pulses were delivered just prior to the presentation of the auditory stimulus such that the last pulse occurred 200 ms before the word onset. 700 ms after spoken word onset, the two target images appeared simultaneously for 500 ms on screen to the left and right of fixation. Subjects were instructed to make their response as quickly and accurately as possible; responses were allowed within a period of 1.5 s after image onset. The following trial started after an inter-trial period ranging between 1.0 and 2.5 s. Subjects responded using their left hand, pressing a left arrow key for the image on the left and a right arrow key for the image on the right. Each response key was associated with lip words and tongue words equally often and in a randomized order. Stimuli were presented on a Windows PC running Matlab 2012 and the Cogent 2000 toolbox (<http://www.vislab.ucl.ac.uk/cogent.php>). The script triggered the onset of the TMS pulses through a direct BNC cable to the TMS device.

Twelve separate right-handed subjects participated in a behavioral control study. They completed four blocks of 70 trials each with the same items and randomization orders as in the TMS experiment but without TMS application and all in one single session.

Data Analysis

Only data from experimental items were analyzed (see Stimuli). A total of 10 trials were excluded because no TMS pulse was given on these trials due to technical problems with the neuronavigation system. Reaction times below 350 ms were excluded as outliers (1.6% of data) as well as those exceeding ± 2 SDs of each subject-and-session-specific mean (4.7% of data). Furthermore, there were five word pairs for which accuracy across all subjects was not significantly different from chance (tested using chi-square tests) and therefore those word pairs were excluded from analysis. The same procedures were applied to the data from the control experiment without TMS.

For the TMS data, we conducted a repeated-measures analysis of variance (ANOVA) with a 2×2 design with factors TMS Location (Lip/Tongue) and Word Type (Lip/Tongue). We analyzed both reaction time (correct responses only) and accuracy (trials where no response was given were discarded, 4.1% of experimental trials) in separate ANOVAs; both were z-score-transformed to each subjects' mean and SD. We also analyzed reaction time data with a linear mixed-effects model using R version 3.0.1 and the package lme4 (Bates et al. 2014). In this model, we employed the same 2×2 design as in the ANOVA but also added random intercepts for each word and each subject. To test significance of effects, we calculated degrees of freedom using the approximation described by Kenward and Roger (1997).

Results

The ANOVA on reaction times revealed a significant main effect for Word Type (lip-word-initial, "lip word" vs. tongue-sound-initial, "tongue word") ($F_{1,12} = 9.6$, $P = 0.009$) and a significant interaction between Word Type and TMS Location ($F_{1,12} = 7.1$, $P = 0.021$). Post-hoc Newman-Keuls tests showed a significant difference for tongue words between the two different stimulation locations ($P = 0.042$), and for the lip TMS Location, a significant

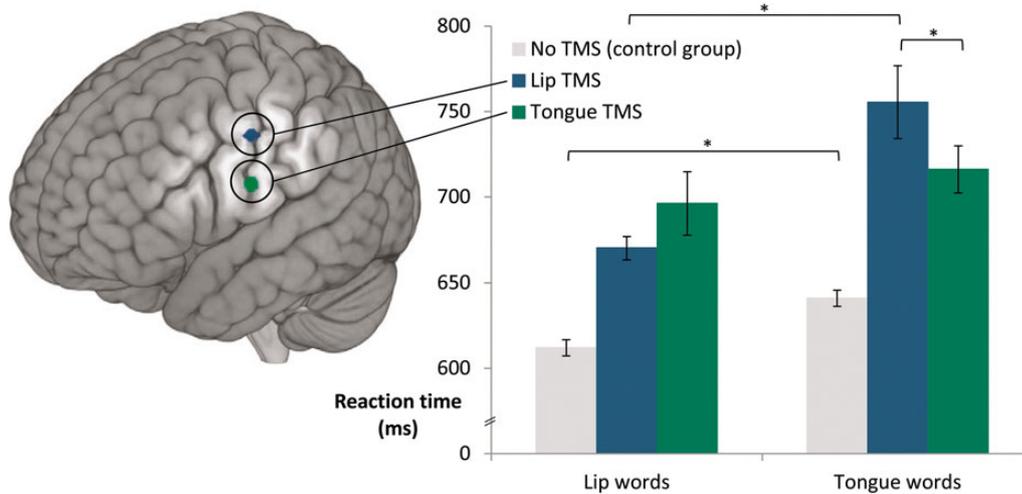


Figure 1. (Left) average stimulation locations for lip and tongue representation shown on a standard MNI brain (lip $x, y, z = -55.4, -9.2, 43.9$; tongue $x, y, z = -59.4, -7.4, 22.8$). (Right) significant interaction of word type by TMS location (reaction time data). The label 'Lip words' denotes words starting with bilabial lip-related phonemes, whereas 'Tongue words' denotes words starting with alveolar tongue-related phonemes. Error bars show ± 1 SEM after removing between-subject variance (Morey 2008) $^*P < 0.05$.

difference between Word Types ($P = 0.001$) (Fig. 1). The linear mixed-effects model confirmed the significant main effect of Word Type ($F = 4.8, P = 0.036$) and significant interaction of Word Type by TMS Location ($F = 5.19, P = 0.023$).

The ANOVA on accuracy data revealed a significant main effect of word type ($F_{1,12} = 5.4, P = 0.038$). Accuracy overall was higher for lip words ($M = 80.6\%$ correct) than that for tongue words ($M = 73\%$ correct) (SE of the difference = 2%). The interaction between Word Type and TMS Location was not significant in the accuracy data ($F_{1,12} = 1.2, P = 0.29$). Nonetheless, there was a similar trend toward lower accuracy with TMS delivered to articulatory areas incongruent with the word-initial phoneme both for lip words (mean difference = 2.2%, SE = 2.6%) and for tongue words (mean difference = 2.3%, SE = 3.5%).

For the behavioral control experiment performed in 12 separate subjects without TMS, a paired t -test revealed that lip words were responded to faster ($M = 612$ ms) than tongue words ($M = 641$ ms), ($P = 0.002$, SE of the difference = 5.1 ms). A paired t -test on the accuracy values revealed no significant difference between word types (lip words: $M = 79\%$ correct, tongue words: $M = 81\%$ correct, $P = 0.55$).

Discussion

We examined the effect of TMS to motor cortex on the comprehension of single spoken words by stimulating the articulatory motor representations of the tongue and lips and measuring word-to-picture-matching responses to words starting with tongue- and lip-produced stop consonants ("lip words" and "tongue words"). The analysis of reaction times showed a significant interaction between word type (lip vs. tongue word) and the location of TMS to motor cortex (lip vs. tongue locus) (Fig. 1). To our knowledge, this result demonstrates for the first time that articulatory motor cortex can exert a causal effect on the comprehension of single meaningful spoken words.

Further analysis of the interaction through post-hoc comparisons showed that reaction times to tongue words were significantly prolonged, by 39 ms in the average, with incongruent TMS stimulation (i.e., to the lip representation) compared with congruent TMS stimulation (of the tongue representation). Correspondingly, reaction times for lip words were delayed by

an average 26 ms with incongruent TMS stimulation relative to congruent stimulation; although this latter post-hoc comparison did not reach significance, both incongruency effects together are manifest in the significant cross-over interaction.

One may ask why TMS produced a clear and significant effect on tongue-sound-initial words but not on lip-sound-initial word processing. Possible explanations include a ceiling effect for lip words, whose RTs were lower than those to tongue words (main effect $P = 0.009$), which was also confirmed in the control experiment without TMS ($P = 0.002$). This indicates that, in spite of all measures taken to control between stimulus materials, tongue-sound-initial items were already more difficult to process, thus being more sensitive to minimal TMS-elicited interference. A further possible explanation can be built on a recent suggestion by Bartoli and colleagues (2015): reconstructing tongue configurations from an acoustic signal might be computationally more demanding due to higher biomechanical complexity and higher degrees of freedom for tongue than for lip movements. This observation might explain why functional TMS interference was easier for tongue- than for lip-sound-initial words. A related point is that lip movements typically engage a smaller part of the motor system than even minimal tongue movements (see Fig. 1 in Pulvermüller et al. 2006). Thus, motor circuits for bilabial phonemes may be more focal than those for alveolar phonemes, making the former easier to functionally influence with TMS. Hence, TMS to focal lip areas might have been able to cause stronger interference effects on the processing of incongruent words than TMS to the relatively distributed tongue representations.

In essence, several reasons may explain why causal effects of motor cortex stimulation were pronounced for tongue-sound-initial words, but marginal for lip words considered separately; importantly, however, the significant interaction of stimulation site and phonemic word type proves that motor cortex stimulation altered word comprehension processes in a phoneme-specific manner. The behavioral pattern revealed by response time results was corroborated by accuracy data (see Results), which showed a relative processing disadvantage for word comprehension when incongruous motor cortex was stimulated. Note that both increase of response times and reduction of accuracies are consistent with a reduction in processing efficacy. Although the error data in themselves did not reach significance,

this consistency is of the essence because it rules out the possibility of a speed-accuracy tradeoff.

The causal role of motor cortex in speech comprehension shown by our results is in line with neurobiological models of language based on Hebbian correlation learning (Braitenberg and Pulvermüller 1992; Pulvermüller et al. 2006; Pulvermüller and Fadiga 2010). When speaking, a speaker usually perceives the self-produced acoustic speech signals, thus implying correlated neuronal activity in inferior-frontal and articulatory motor cortex and in superior-temporal auditory areas. Correlation learning commands that frontotemporal circuits important for both language production and comprehension emerge. That temporal areas, apart from fronto-central ones, play a role in speech production appears uncontroversial (Lichtheim 1885; Paus et al. 1996). The present study now presents unambiguous evidence that, apart from superior-temporal areas, also fronto-central sites and especially motor cortices take a role in speech comprehension, thus supporting the action-perception model.

We hypothesized that stimulation of a congruent sector of articulatory motor cortex (lip or tongue) would lead to faster word comprehension than the stimulation of incongruous motor areas, where competing motor programs may be activated thus causing a degree of interference. This hypothesis was confirmed by the significant interaction of stimulation site by word type. Our baseline condition without TMS was performed with a different population of subjects so that the faster responses without TMS may be due to general or specific effects of TMS or rather to a difference between subject groups. We did not include a sham TMS condition for the subjects receiving TMS because this would have made stimulus repetition unavoidable. Therefore, we cannot determine with certainty, based on the present data, whether the effects are indeed a result of interference and whether facilitation may have played a role. The previous TMS study showing motor cortex influence on phoneme classification indeed found evidence for both interference and facilitations (D'Ausilio et al. 2009). This issue should be further investigated for word comprehension. Furthermore, the observed interaction effect of word type and stimulation site in single word comprehension could have been caused directly by the stimulation of primary motor cortex neurons or, as D'Ausilio et al. (2009) suggested, by an indirect effect of primary on premotor circuits, which, in this view, would influence comprehension. Still, our results show clearly that superior-temporal cortex is not the only area causally involved in speech perception and comprehension.

Earlier studies had already shown that TMS to articulatory motor or premotor cortex has an effect on speech classification performance (Meister et al. 2007; D'Ausilio et al. 2009; Möttönen and Watkins 2009; Sato et al. 2009). However, this earlier work has been criticized because explicit phoneme classification is never necessary in linguistic communication, and effects could have emerged at the level of phoneme classification, rather than perception per se (Hickok 2009; McGettigan et al. 2010). This interpretation can now be ruled out on the basis of recent studies combining TMS and EEG/MEG by Möttönen et al. (2013, 2014), who found that after repetitive TMS (rTMS) of the lip representation in motor cortex, the mismatch negativity (MMN) to speech sounds was reduced even when subjects passively perceived these sounds without explicit classification tasks. Other studies could show that speech listening automatically induces changes in articulatory movements, thus providing further arguments for causal functional links between the perception and articulation of speech sounds in the absence of classification tasks (Yuen et al. 2010; D'Ausilio et al. 2014). Nonetheless, these striking results still do not speak to the comprehension issue,

that is, to the question of whether earlier results on the relevance of motor systems generalize to the crucial level of semantic comprehension of meaningful speech. In order to monitor comprehension, a task needs to involve at least one of the critical aspects of semantics, for example the relationship between words and objects, as is the case in the word-to-picture-matching task used most frequently in psycho- and neuro linguistics to investigate semantic understanding.

Placing the phonemes in the context of whole words allowed us to investigate the role of motor cortex sectors related to these phonemes when these serve their normal function as meaning-discriminating units. A recent study by Krieger-Redwood et al. (2013) used a similar approach and tried to disentangle speech perception and comprehension processes. However, these authors used classification tasks, which in our view makes their results difficult to interpret. Subjects had to explicitly categorize either the final speech sound of words into phonological categories (e.g., 'k' vs. 't') or their meaning into semantic categories (e.g., large vs. small). The authors found that rTMS to premotor cortex only delayed reaction times for phonological but not semantic judgments and conclude that the role of premotor cortex does not extend to accessing meaning in speech comprehension. Furthermore, that study used to a degree artificial 'cross-spliced' stimuli and TMS was to relatively dorsal premotor cortex (MNI $x, y, z = -52.67, -6.67, 43$), which is consistent with our present lip site but distant from our tongue site, so that only a fraction of tongue-related phonemes might be affected. Finally, as mentioned earlier, both the phoneme classification and semantic classification tasks require cognitive processes of comparison and classification going beyond speech perception and comprehension. Therefore, it is unclear whether the delays in phonological categorization responses originate at the level of phoneme perception or rather at those of comparison and categorization. The delays in phonological and their absence in semantic categorization reported by these authors may therefore speak more to the categorization issue than to that of speech perception and understanding. In contrast, our results provide decisive information to resolve the speech comprehension debate as they show a causal role of articulatory motor cortex on speech comprehension when words serve their normal function as carriers of meaning.

Earlier TMS studies showing motor cortex influences on speech perception have also been criticized because of noise overlay of the auditory stimuli, which, apart from the use of 'artificial' phonological tasks, contributed to unnatural listening conditions. Though it can be argued that speech perception without noise is actually the exception rather than the norm in everyday life (D'Ausilio et al. 2012), it has been claimed that motor influences on speech perception might disappear when stimuli are not overlaid with noise (Rogalsky et al. 2011). Our results argue against this view by showing motor cortex influences on speech comprehension using non-synthesized stimuli without noise overlay. Although the reduction of sound pressure level (SPL) constitutes a decrease in signal-to-noise ratio (SNR), we see a major difference between our present approach and earlier studies which used noise to mask speech stimuli. By reducing SPL, the critical variable of change was the *degree of redundancy of the speech signal* (for discussion, see Wilson 2009). For example, it is enough to hear part of the vowel [u] to identify the entire lexical item "pool", because co-articulation effects provide the listener with information about the preceding and subsequent phonemes (see, e.g., Warren and Marslen-Wilson 1987). By minimizing redundancy through SNR reduction we observed a motor influence on the comprehension of spoken words not masked by

noise, in line with the observation that earlier TMS studies (Fadiga et al. 2002; Watkins et al. 2003) and fMRI studies (Pulvermüller et al. 2006; Grabski et al. 2013) also found motor cortex activation in speech perception using stimuli without noise overlay. Furthermore, a recent study (Bartoli et al. 2015) showed that TMS to motor cortex causes interference in a speech discrimination task where task difficulty is induced by inter-speaker variability of naturally spoken syllables without noise.

Finally, our results also address two further objections that had been raised against earlier TMS studies investigating phoneme discrimination. First, it has been argued that motor cortex involvement in speech perception only occurs in tasks requiring phonemic segmentation (Sato et al. 2009), although the null effects in other tasks reported by that study could be due to differences in task complexity. Secondly, it is claimed that TMS to motor systems does not modulate perception of speech sounds but rather secondary decision processes such as response bias (Venezia et al. 2012). However, the fact that our task did not require segmenting or classifying phonemes argues against these views. Recognizing the word-initial sounds as either a bilabial or alveolar stop consonant was not the task per se but rather was implicit to understanding the whole spoken word and mapping it onto its meaning, as it is typical for natural language processing. Our results thus support that TMS to motor cortex indeed affects speech perception and comprehension (rather than any possible response bias or segmentation process). These conclusions are also in line with two recent studies which found that the TMS to motor cortex impairs syllable discrimination in a task unaffected by response bias (Rogers et al. 2014) and that rTMS to motor cortex affects sensitivity of speech discrimination, but not response bias (Smalle et al. 2015).

In summary, we here show a causal influence of articulatory motor cortex on the comprehension of meaningful words in the standard psycholinguistic task of word-to-picture matching. Furthermore, our results show that such effects can be obtained with naturally spoken stimuli without artificial noise overlay and are not due to response bias or other features epiphenomenal to the comprehension process. In the wider neuroscience debate about the frontal cortex' role in perceptual processing, the observed causal effect of motor cortex activation on language comprehension now demonstrates that the human motor system is not just activated in perception and comprehension, as previous research amply demonstrated, but that it also serves a critical role in the comprehension process itself. These results support an action-perception model of language (Pulvermüller and Fadiga 2010) and are inconsistent with classic modular accounts attributing speech comprehension exclusively to temporal areas and denying a general causal contribution of motor systems to language comprehension (e.g., Wernicke 1874; Lichtheim 1885; Geschwind 1970; Hickok 2009, 2014).

Supplementary Material

Supplementary Material can be found at <http://www.cercor.oxfordjournals.org/online>.

Funding

Funding for this research was provided by the Deutsche Forschungsgemeinschaft (DFG, Excellence Cluster 'Languages of Emotion', Research Grants Pu 97/15-1, Pu 97/16-1 to F.P.), the Berlin School of Mind and Brain (doctoral scholarship to M.R.S.), the Engineering and Physical Sciences Research Council and Behavioural and Brain Sciences Research Council (BABEL grant,

EP/J004561/1 to F.P.) and the Freie Universität Berlin. Funding to pay the Open Access publication charges for this article was also provided by the above.

Notes

Conflict of Interest: None declared.

References

- Alho J, Lin F-H, Sato M, Tiitinen H, Sams M, Jääskeläinen IP. 2014. Enhanced neural synchrony between left auditory and pre-motor cortex is associated with successful phonetic categorization. *Front Psychol.* 5:394.
- Bak TH, O'Donovan DG, Xuereb JH, Boniface S, Hodges JR. 2001. Selective impairment of verb processing associated with pathological changes in Brodmann areas 44 and 45 in the motor neurone disease-dementia-aphasia syndrome. *Brain.* 124:103-120.
- Bartoli E, D'Ausilio A, Berry J, Badino L, Bever T, Fadiga L. 2015. Listener-speaker perceived distance predicts the degree of motor contribution to speech perception. *Cereb Cortex.* 25:281-288.
- Basso A, Casati G, Vignolo LA. 1977. Phonemic identification defect in aphasia. *Cortex.* 13:85-95.
- Bates D, Maechler M, Bolker B, Walker S. 2014. lme4: linear mixed-effects models using Eigen and S4. R package version 1.1-6. Available from: <http://CRAN.R-project.org/package=lme4>.
- Blumstein SE, Baker E, Goodglass H. 1977. Phonological factors in auditory comprehension in aphasia. *Neuropsychologia.* 15:19-30.
- Blumstein SE, Burton M, Baum S, Waldstein R, Katz D. 1994. The role of lexical status on the phonetic categorization of speech in aphasia. *Brain Lang.* 46:181-197.
- Bookheimer S. 2002. Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annu Rev Neurosci.* 25:151-188.
- Boulenger V, Mechtouff L, Thobois S, Broussolle E, Jeannerod M, Nazir TA. 2008. Word processing in Parkinson's disease is impaired for action verbs but not for concrete nouns. *Neuropsychologia.* 46:743-756.
- Braitenberg V, Pulvermüller F. 1992. Entwurf einer neurologischen Theorie der Sprache. *Naturwissenschaften.* 79:103-117.
- Caplan D, Gow D, Makris N. 1995. Analysis of lesions by MRI in stroke patients with acoustic-phonetic processing deficits. *Neurology.* 45:293-298.
- Chevillet MA, Jiang X, Rauschecker JP, Riesenhuber M. 2013. Automatic phoneme category selectivity in the dorsal auditory stream. *J Neurosci.* 33:5208-5215.
- D'Ausilio A, Bufalari I, Salmas P, Fadiga L. 2012. The role of the motor system in discriminating normal and degraded speech sounds. *Cortex.* 48:882-887.
- D'Ausilio A, Maffongelli L, Bartoli E, Campanella M, Ferrari E, Berry J, Fadiga L. 2014. Listening to speech recruits specific tongue motor synergies as revealed by transcranial magnetic stimulation and tissue-Doppler ultrasound imaging. *Philos Trans R Soc Lond B Biol Sci.* 369:20130418.
- D'Ausilio A, Pulvermüller F, Salmas P, Bufalari I, Begliomini C, Fadiga L. 2009. The motor somatotopy of speech perception. *Curr Biol.* 19:381-385.
- Du Y, Buchsbaum BR, Grady CL, Alain C. 2014. Noise differentially impacts phoneme representations in the auditory and speech motor systems. *Proc Natl Acad Sci.* 111:7126-7131.

- 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 CA, Turvey MT. 2006. The motor theory of speech perception reviewed. *Psychon Bull Rev*. 13:361–377.
- Geschwind N. 1970. The organization of language and the brain. *Science*. 170:940–944.
- Gough PM, Nobre AC, Devlin JT. 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, Baciú M, Le Bas J-F, Sato M. 2013. Shared and distinct neural correlates of vowel perception and production. *J Neurolinguistics*. 26:384–408.
- Heister J, Würzner K-M, Bubenze J, Pohl E, Hanneforth T, Geyken A, Kliegl R. 2011. dlexDB—eine lexikalische Datenbank für die psychologische und linguistische Forschung. *Psychol Rundsch*. 62:10–20.
- Hickok G. 2009. Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *J Cogn Neurosci*. 21:1229–1243.
- Hickok G. 2014. The myth of mirror neurons: the real neuroscience of communication and cognition. WW Norton & Company.
- Kammer T, Baumann LW. 2010. Phosphene thresholds evoked with single and double TMS pulses. *Clin Neurophysiol*. 121:376–379.
- Kemmerer D, Rudrauf D, Manzel K, Tranel D. 2012. Behavioral patterns and lesion sites associated with impaired processing of lexical and conceptual knowledge of actions. *Cortex*. 48:826–848.
- Kenward MG, Roger JH. 1997. Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*. 53:983–997.
- Komeilipoor N, Vicario CM, Daffertshofer A, Cesari P. 2014. Talking hands: tongue motor excitability during observation of hand gestures associated with words. *Front Hum Neurosci*. 8:767.
- Krieger-Redwood K, Gaskell MG, Lindsay S, Jefferies B. 2013. The selective role of premotor cortex in speech perception: a contribution to phoneme judgements but not speech comprehension. *J Cogn Neurosci*. 25:2179–2188.
- Lee YS, Turkeltaub P, Granger R, Raizada RDS. 2012. Categorical speech processing in Broca's area: an fMRI study using multivariate pattern-based analysis. *J Neurosci*. 32:3942–3948.
- Lieberman AM, Mattingly IG. 1985. The motor theory of speech perception revised. *Cognition*. 21:1–36.
- Lichtheim L. 1885. On aphasia. *Brain*. 7:433–484.
- Liebenthal E, Sabri M, Beardsley SA, Mangalathu-Arumana J, Desai A. 2013. Neural dynamics of phonological processing in the dorsal auditory stream. *J Neurosci*. 33:15414–15424.
- Mahon BZ, Caramazza A. 2008. A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *J Physiol Paris*. 102:59–70.
- McGettigan C, Agnew ZK, Scott SK. 2010. Are articulatory commands automatically and involuntarily activated during speech perception? *Proc Natl Acad Sci USA*. 107:E42.
- Meister IG, Wilson SM, Deblieck C, Wu AD, Iacoboni M. 2007. The essential role of premotor cortex in speech perception. *Curr Biol*. 17:1692–1696.
- Miceli G, Gainotti G, Caltagirone C, Masullo C. 1980. Some aspects of phonological impairment in aphasia. *Brain Lang*. 11:159–169.
- Moineau S, Dronkers NF, Bates E. 2005. Exploring the processing continuum of single-word comprehension in aphasia. *J Speech Lang Hear Res*. 48:884.
- Moliadze V, Zhao Y, Eysel U, Funke K. 2003. Effect of transcranial magnetic stimulation on single-unit activity in the cat primary visual cortex. *J Physiol*. 553:665–679.
- Morey RD. 2008. Confidence intervals from normalized data: a correction to Cousineau (2005). *Tutor Quant Methods Psychol*. 4:61–64.
- Mottaghy FM, Sparing R, Töpper R. 2006. Enhancing picture naming with transcranial magnetic stimulation. *Behav Neurol*. 17:177–186.
- Möttönen R, Dutton R, Watkins KE. 2013. Auditory-motor processing of speech sounds. *Cereb Cortex*. 23:1190–1197.
- Möttönen R, van de Ven GM, Watkins KE. 2014. Attention fine-tunes auditory-motor processing of speech sounds. *J Neurosci*. 34:4064–4069.
- Möttönen R, Watkins KE. 2009. Motor representations of articulators contribute to categorical perception of speech sounds. *J Neurosci*. 29:9819–9825.
- Möttönen R, Watkins KE. 2012. Using TMS to study the role of the articulatory motor system in speech perception. *Aphasiology*. 26:1103–1118.
- Murakami T, Ugawa Y, Ziemann U. 2013. Utility of TMS to understand the neurobiology of speech. *Front Psychol*. 4:446.
- Neininger B, Pulvermüller F. 2001. The right hemisphere's role in action word processing: a double case study. *Neurocase*. 7:303–317.
- Neininger B, Pulvermüller F. 2003. Word-category specific deficits after lesions in the right hemisphere. *Neuropsychologia*. 41:53–70.
- Niyazov DM, Butler AJ, Kadah YM, Epstein CM, Hu XP. 2005. Functional magnetic resonance imaging and transcranial magnetic stimulation: effects of motor imagery, movement and coil orientation. *Clin Neurophysiol*. 116:1601–1610.
- Oldfield RC. 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*. 9:97–113.
- Osnes B, Hugdahl K, Specht K. 2011. Effective connectivity analysis demonstrates involvement of premotor cortex during speech perception. *Neuroimage*. 54:2437–2445.
- Paus T, Pery DW, Zatorre RJ, Worsley KJ, Evans AC. 1996. Modulation of cerebral blood flow in the human auditory cortex during speech: role of motor-to-sensory discharges. *Eur J Neurosci*. 8:2236–2246.
- Pulvermüller F, Fadiga L. 2010. Active perception: sensorimotor circuits as a cortical basis for language. *Nat Rev Neurosci*. 11:351–360.
- Pulvermüller F, Huss M, Kherif F, Moscoso Del Prado Martin F, Hauk O, Shtyrov Y. 2006. Motor cortex maps articulatory features of speech sounds. *Proc Natl Acad Sci USA*. 103:7865–7870.
- Pulvermüller F, Shtyrov Y, Ilmoniemi R. 2003. Spatiotemporal dynamics of neural language processing: an MEG study using minimum-norm current estimates. *Neuroimage*. 20:1020–1025.
- Rilling JK, Glasser MF, Preuss TM, Ma X, Zhao T, Hu X, Behrens TE. 2008. The evolution of the arcuate fasciculus revealed with comparative DTI. *Nat Neurosci*. 11:426–428.
- Rizzolatti G, Cattaneo L, Fabbri-Destro M, Rozzi S. 2014. Cortical mechanisms underlying the organization of goal-directed actions and mirror neuron-based action understanding. *Physiol Rev*. 94:655–706.
- Rizzolatti G, Fadiga L, Gallese V, Fogassi L. 1996. Premotor cortex and the recognition of motor actions. *Cogn Brain Res*. 3:131–141.
- Rogalsky C, Love T, Driscoll D, Anderson SW, Hickok G. 2011. Are mirror neurons the basis of speech perception? Evidence from five cases with damage to the purported human mirror system. *Neurocase*. 17:178–187.

- Rogers JC, Möttönen R, Boyles R, Watkins KE. 2014. Discrimination of speech and non-speech sounds following theta-burst stimulation of the motor cortex. *Front Psychol.* 5:754.
- Rossini PM. 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. *Electroenceph Clin Neurophysiol.* 91:79–92.
- Sato M, Tremblay P, Gracco VL. 2009. A mediating role of the premotor cortex in phoneme segmentation. *Brain Lang.* 111:1–7.
- Smalle EHM, Rogers J, Möttönen R. 2015. Dissociating contributions of the motor cortex to speech perception and response bias by using transcranial magnetic stimulation. *Cereb Cortex.* 25:3690–3698.
- Utman JA, Blumstein SE, Sullivan K. 2001. Mapping from sound to meaning: reduced lexical activation in Broca's aphasics. *Brain Lang.* 79:444–472.
- Van Casteren M, Davis MH. 2006. Mix, a program for pseudo-randomization. *Behav Res Methods.* 38:584–589.
- Venezia JH, Saberi K, Chubb C, Hickok G. 2012. Response bias modulates the speech motor system during syllable discrimination. *Front Psychol.* 3:157.
- Walsh V, Cowey A. 2000. Transcranial magnetic stimulation and cognitive neuroscience. *Nat Rev Neurosci.* 1:73–80.
- Warren P, Marslen-Wilson W. 1987. Continuous uptake of acoustic cues in spoken word recognition. *Percept Psychophys.* 41:262–275.
- Watkins KE, Strafella AP, Paus T. 2003. Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia.* 41:989–994.
- Wernicke C. 1874. *Der Aphasische Symptomenkomplex. Eine psychologische Studie auf anatomischer Basis.* Breslau, Germany: Kohn und Weigert.
- Wilson SM. 2009. Speech perception when the motor system is compromised. *Trends Cogn Sci.* 13:329–330.
- Wilson SM, Saygin AP, Sereno MI, Iacobini M. 2004. Listening to speech activates motor areas involved in speech production. *Nat Neurosci.* 7:701–702.
- Yee E, Blumstein SE, Sedivy JC. 2008. Lexical-semantic activation in Broca's and Wernicke's aphasia: Evidence from eye movements. *J Cogn Neurosci.* 20:592–612.
- Yuen I, Davis MH, Brysbaert M, Rastle K. 2010. Activation of articulatory information in speech perception. *Proc Natl Acad Sci USA.* 107:592–597.