

## Sensorimotor and Motorsensory Interactions in Speech

Vincent L Gracco<sup>1,2,3</sup>, Ekaterina Klepousniotou<sup>4</sup>, Inbal Itzhal<sup>1,2</sup>, Shari R. Baum<sup>1,2</sup>

<sup>1</sup>Centre for Research on Language, Mind and Brain, McGill University, Montreal

<sup>2</sup>School of Communication Sciences & Disorders, McGill University, Montreal

<sup>3</sup>Haskins Laboratories, New Haven, CT USA

<sup>4</sup>University of Leeds, Leeds, UK

E-mail: [vincent.gracco@mcgill.ca](mailto:vincent.gracco@mcgill.ca)

### Abstract

*A long-standing issue in psycholinguistics is whether language production and language comprehension share a common neural substrate. Recent neuroimaging studies of speech appear to support overlap of brain regions for both production and perception. However, what is not known is how to interpret the perceptual activation of motor regions. In the following, the brain regions associated with producing heard speech are described to identify the sensorimotor components of the speech motor network. The brain regions associated with speech production are then examined for their activation during passive perception of lexical items presented as heard words, pictures and printed text. A number of overlapping cortical and subcortical areas were activated during both perception and production. Interestingly, all brain areas associated with passive perception increased their activation for speech production. The increased activation in the classical sensory/perceptual areas for production suggests an interactive process in which motor areas project back to sensory/perceptual areas reflecting a binding of perception (sensory) and production (motor) regions within the network.*

### 1 Introduction

Speech refers to the processes that are used in the production and perception of sounds in spoken language. The idea that these two processes might be linked was first detailed in the motor theory of speech perception (MTSP) [1]. The MTSP was an extension of earlier theories of perception in which motor actions were viewed as integral components of the perceptual process [2]. With the discovery of the mirror neuron system [3] an apparent neurobiological basis for production-perception overlap was identified and since then neuroimaging

evidence has been accumulating demonstrating that overlapping brain areas are activated for both speech production and speech perception. Brain areas involved in the planning and execution of speech gestures (notably the left inferior frontal gyrus, the premotor and primary motor cortex) and areas subserving proprioception related to mouth movements (somatosensory cortex) have been found to be activated during both auditory speech processing, visual speech processing and congruent/incongruent audiovisual speech perception [4-9]. Moreover, repetitive and single pulse TMS applied over frontal motor areas has been shown to interact with speech processing and even impact speech perception indicating that the motor system may be a functional component in speech perceptual processes [10-12] under certain processing conditions [13]. One of the implications from neuroimaging studies of speech perception is that observation of action and even action itself is not necessary for activating speech motor areas. It has been observed that merely listening to auditory sounds (speech and nonspeech) and hearing sounds (or the actions associated with sounds) activate the motor system [5, 10, 14].

In contrast, speech production has long been considered a sensorimotor process. The functional sensorimotor nature of speech production has been demonstrated in a number of studies employing manipulations of both somatosensory and auditory feedback. Introducing unexpected dynamic mechanical loads to some portion of the speech production system results in rapid (on-line), compensatory changes in all active muscles and movements [15-18]. Similarly, manipulating a subject's own auditory feedback changes speech motor output in characteristic ways to offset the sensory modification [19]. Sensory-based adjustments in speech motor output have also been used to investigate adaptive changes following mechanical [20] and auditory [21,22] manipulations. In these

studies, sensory feedback is gradually changed over time and speech motor output compensates over time for the sensory manipulation. More recently it has been shown that while auditory information is often assumed as the dominant sensory modality, somatosensory information has significant and substantial impact on speech motor learning when auditory feedback is completely absent [23]. Unexpected perturbations are assumed to tap into the existing sensorimotor organization while adaptive changes are assumed to represent a re-calibration (or re-learning) of the mapping between sensory signals and motor output.

However, in order to fully understand the extent of any perception-production overlap or the functional significance of their interactions, more empirical work is needed. A necessary prerequisite for informed studies of perception-production interactions is a solid understanding and detailed description of the brain areas in which perception-production interactions may take place, and, more importantly, an understanding of the manner in which speech tasks modulate brain function. In what follows, some recent findings that highlight the manner in which sensory perceptual systems interact with speech action will be presented. The motivation is to provide a better understanding of the neural network that is the basis for speech processing in order to appropriately interpret the perception-production link.

### Neural Substrate

A common finding in studies of speech production is that the neural substrate is distributed among multiple cortical and subcortical regions [see 24,25 for reviews] and activated for a range of tasks including isometric rhythmic tongue contraction, syllable repetition, word production, counting, reciting over-learned nursery rhymes, and propositional speech and discourse. Using a functional imaging approach for overt speech production detailed in [26] and the repetition of heard words and sentences, we have identified what appears to be the fundamental sensorimotor network for speech motor production. The distributed network, pictured in the figures below, includes the primary motor and somatosensory cortical areas, lateral and medial premotor areas, the insula, the anterior cingulate cortex, distributed activation along the supratemporal plane of the superior temporal gyrus (STG), the sensorimotor portion of the basal ganglia (putamen, substantia nigra, subthalamic nucleus), the sensorimotor thalamus and the superior and inferior cerebellar hemispheres.

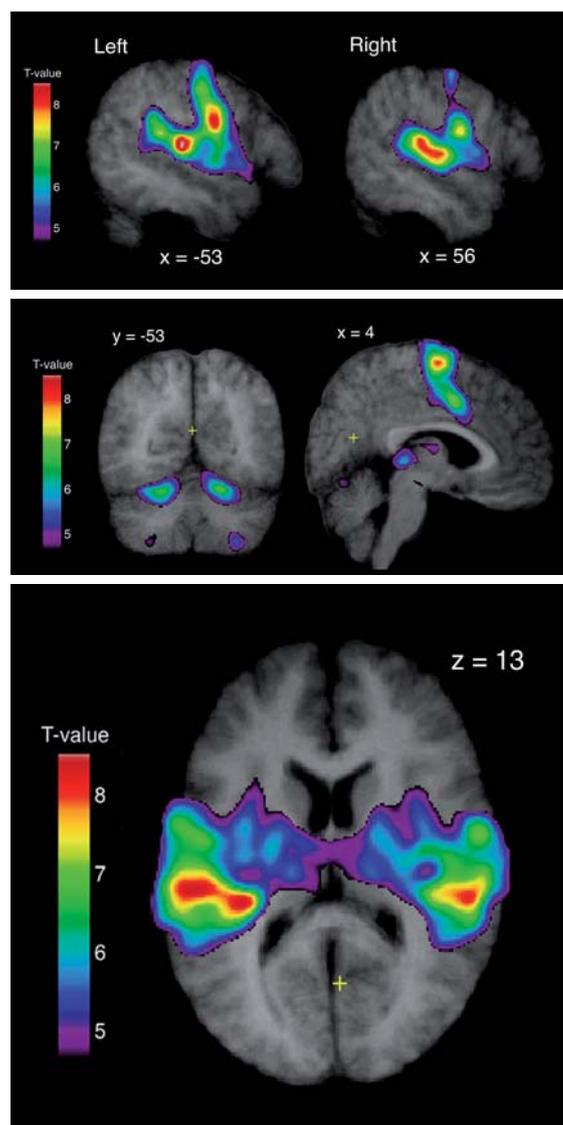


Figure 1. *Statistical parametric maps of brain activation following the repetition of heard speech (words and short sentences). The activation represents the results of a subtraction of passive listening to words and sentences from the experimental condition (listening and repeating heard speech). Top panel is a left and right lateral view of the cortical surface area activation. The middle panel is a coronal (left) and midsagittal view and the bottom panel is an axial section taken through the thalamus and basal ganglia. The stereotaxic coordinates (x, y and z) refer to the left-right, posterior-anterior, and ventral-dorsal (respectively) distance from the anterior commissure.*

In contrast the neural substrate for speech perception, while extensive, involves a more focal network. The dual-stream model of speech processing recently [27] places much of the initial auditory processing in the primary auditory cortices along the supratemporal plane. Subsequent phonological level processing involves the middle and posterior portions of the superior temporal sulcus (STS). At this point, two broad paths diverge

with the dorsal most path (the dorsal stream) representing phonological or sensory representations that are mapped onto speech motor representations. Sensorimotor components of the dorsal stream include the parieto-temporal operculum (pSTG) and more anterior locations on the frontal lobe including the posterior inferior frontal gyrus, the anterior insula and the dorsal premotor cortex on the left hemisphere. According to the Hickok and Poeppel (2007), the dorsal path is involved in translating acoustic speech signals into articulatory representations and is hypothesized to have an auditory-motor integration function.

Observations of motor system activation during speech perception has resulted in two distinct proposals with regard to the functional processes for speech perception and production. From the viewpoint of Hickok and Poeppel (2007) the motor system is not recruited for perception (or comprehension) per se as it is the ventral stream that transforms sound into meaning and the ventral stream does not interface directly with the motor system. In contrast, the dorsal stream directly interfaces with the motor system and performs a coordinate transformation of the heard acoustic properties to articulatory actions for speech production. Perception is viewed as an auditory process and any interaction with the motor system is related directly to implementing production. In contrast, the mirror-neuron perspective has evolved into a motor resonance process in which perception links to action and an interacting network is activated. Recently two forms of resonance have been proposed [28], a communicative motor resonance and a referential motor resonance. The distinction between these two forms of motor resonance is in the specificity of the response. In a communicative motor resonance, the motor system simulates the production of the utterance. If a portion of the motor system is activated that is directly related to the body portion that is reflected in the action, there is referential motor resonance. Fischer and Zwann (2008) suggest that these two types of motor resonance occur simultaneously during language comprehension.

## 2 Methods

Fifty lexical items (all nouns selected from the Snodgrass and Vanderwart set) were presented either in written, picture or auditory form. Three conditions (rest, passive stimulus presentation and overt word production) of ten trials each were blocked with the order of presentation counterbalanced across subjects. Prior to each block subjects were instructed to either passively view or listen to the stimuli or

they were instructed to name the picture, read the word aloud or repeat the auditory word. A rest condition was interspersed and subjects were instructed to lay quietly with eyes open to a blank screen. All stimuli and subject responses were timed to occur during the silent interval (Figure 2). Subjects: included ten right-handed adults (average age=26 years) balanced for gender.

Thirty six 4x4x4 mm axial slices were acquired on the 3 T Siemens Trio scanner at the Montreal Neurological Institute in 2.16 sec with a 2.84 sec silent interval (TR=5 sec). All stimuli were presented and all responses were produced during the silent interval. All speech responses were recorded directly to disk. Each session included two experimental runs and a high-resolution anatomical image.

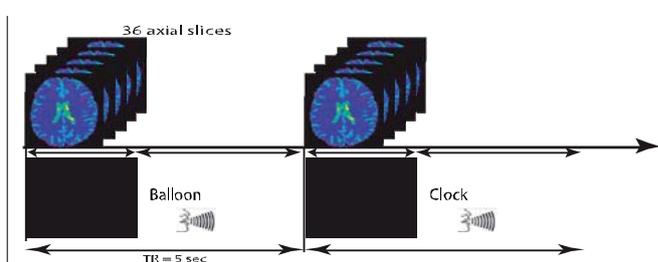


Figure 2. Schematic of the experimental design illustrating the fMRI acquisition and stimulus delivery timing.

After motion correction, fMRI data were low pass filtered, normalized and statistical analysis was performed using a linear model with correlated errors (NeuroLens). Statistical parametric maps were obtained for each condition and subject and then for the group. In addition, regions of interest (8 mm volumes) were selected from all activated regions for all conditions (passive and active) and extracted and averaged for each subject. The ROI were converted to a percent change relative to the rest baseline.

## 3 Results

Passive presentation of lexical items, regardless of stimulus modality, resulted in activation in a number of motor and sensorimotor areas including primary motor cortex, premotor cortex (left side), anterior insula ( $l > r$ ) and SMA.

Presented in Figures 4-6 are some of the ROI results illustrating the activation of different cortical (motor cortex, anterior insula) and subcortical (putamen) areas for the passive and active conditions. The conditions are picture viewing (Picture), picture naming (Name), printed text viewing (Print), reading aloud (Read), listening to auditory words (Auditory) and repeating heard words (Repeat).

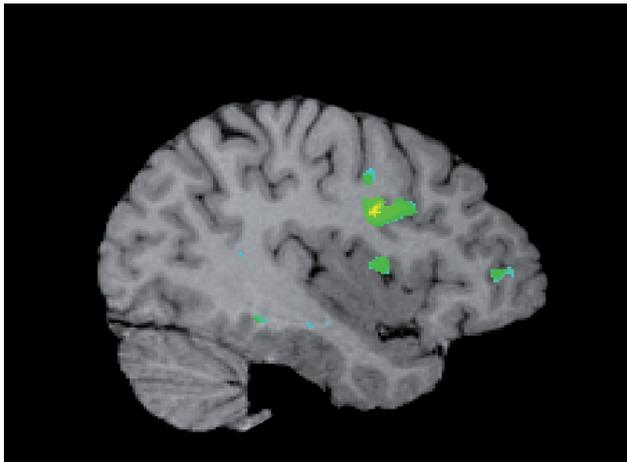


Figure 3. Sagittal section showing left hemisphere activation of the premotor cortex and anterior-medial insula for the passive condition. The activation is a result of a conjunction analysis of all three passive conditions.

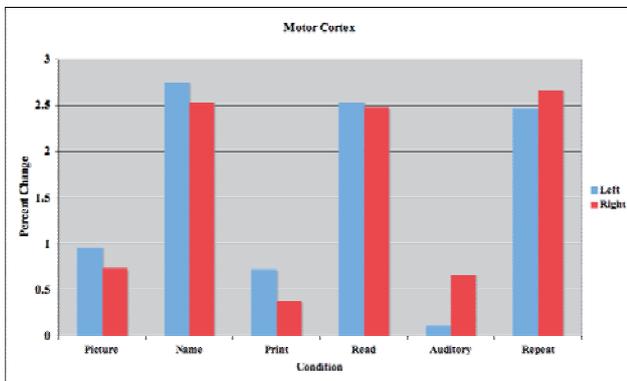


Figure 4. Motor cortex activation.

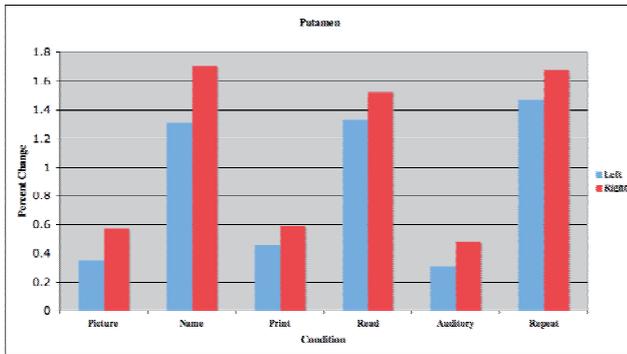


Figure 5. Putamen activation.

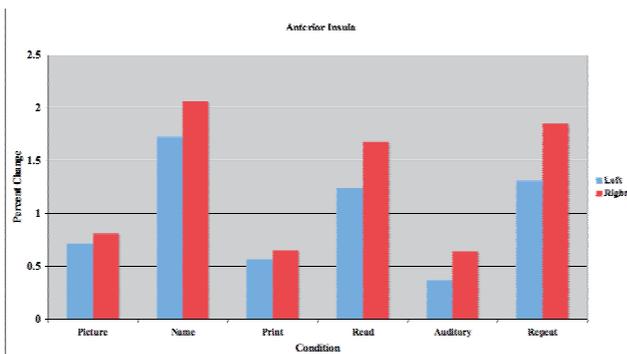


Figure 6. Insular activation.

As can be seen, there is greater activation in these areas for the active (production) condition compared to the passive condition. Moreover, in many instances there is greater passive activation for picture and text viewing compared to listening to words. While the active-passive differences (active > passive) in motor (or sensorimotor) areas are expected, cortical areas that are usually thought of as primarily perceptual also showed the same changes.

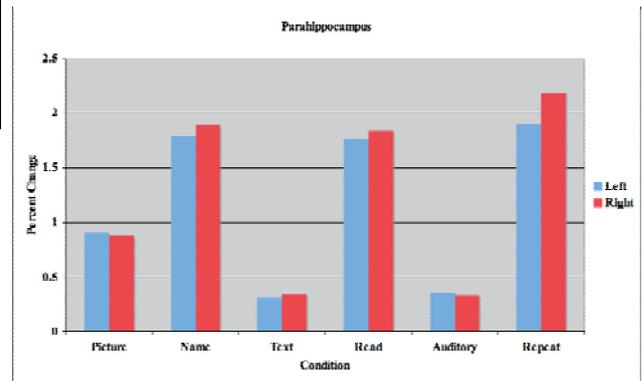


Figure 7. Parahippocampus activation.

Shown in Figure 7 is activation in the left and right parahippocampus, a region that plays a role in memory encoding and retrieval and in recognition of scenes. What can be seen is activation for the picture viewing, and reduced activation for the text viewing and word listening. However, all production conditions resulted in increased activation in this region. This was a common observation in a number of occipital and temporal regions.

### 4 Conclusions

Overall the data reflect motor area activation following passive presentation of the same lexical items presented through different representations. In all cases, the motor areas activated for speech production (primary motor cortex, supplementary motor area, the insula, lateral premotor cortex and the putamen) demonstrated some degree of activation during the passive perceptual tasks. In contrast, and not shown in the present analysis, there was no activation in the cerebellar motor region during any of the passive tasks. Also, the posterior STG, area SPT suggested by [27] to play a role in acoustic-articulatory transformation, was activated but only for the listening condition, not the passive word or picture condition. It should also be noted, that the evidence of motor area activation during passive viewing/listening in the whole brain analyses was only observed following a conjunction of all three conditions, and in order to find evidence of motor area activity in the separate passive conditions, ROI extraction was required. This

suggests that the level of activation, and hence the contribution of perceptual input onto motor areas, is by no means substantial. This does not rule out a contribution of motor areas to perceptual processes but suggests that either motor activation plays a minimal role or an active task is required to enhance the motor area contribution.

In contrast to previous studies demonstrating motor area activation during action perception, we found evidence of activation in the presence of simple concrete nouns. Moreover, in contrast to previous studies demonstrating passive auditory related activation, we found that passive visual input often resulted in more activation compared to the auditory condition. The activation of motor areas during passive perception seems to reflect a communicative motor resonance [28] in which lexical items, regardless of the domain of presentation, activates an obligatory set of connections, in which the name of the object is activated.

Finally, the increased activation in sensory/perceptual areas in the speech production conditions suggests that speech motor areas may also project back onto the perceptual areas that are activated by environmental stimuli. This feedback projection may be one way in which sensory input is used to facilitate motor output through increased and oscillatory activation of the overall network. The functional significance of these potential sensorimotor and motorsensory interactions remains to be clarified.

## 5 References

- [1] A.M. Liberman & I.G. Mattingly. The motor theory of speech perception revised. *Cognition*, 21: 1-36, 1985.
- [2] L. Festinger, C.A. Burnham, H. Ono & D. Bamber, D. Efference and the conscious experience of perception. *Journal of Experimental Psychology Monographs*, 74: 1-36, 1967.
- [3] G. Rizzolatti, L. Fadiga, M. Matelli, V. Bettinardi, E. Paulesu, et al. Localization of grasp representation in humans by PET: 1. Observation versus execution. *Experimental Brain Research*, 111:246-52, 1996.
- [4] G.A. Calvert & R. Campbell. Reading speech from still and moving faces: The neural substrates of visible speech. *Journal of Cognitive Neuroscience*, 15(1): 57-70, 2003.
- [5] S.M. Wilson, A.P. Saygin, M.I. Sereno & M. Iacoboni. Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, 7: 701-702, 2004.
- [6] J.I. Skipper, H.C. Nusbaum & S.L. Small. Listening to talking faces: Motor cortical activation during speech perception. *NeuroImage*, 25: 76-89, 2005.
- [7] J.I. Skipper, V. Van Wassenhove, H.C. Nusbaum & S.L. Small. Hearing lips and seeing voices: how cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex*, 17: 2387-2399, 2007.
- [8] F. Pulvermuller, M. Huss, F. Kherif, F. Moscoso del Prado, O. Hauk & Y. Shtyrov. Motor cortex maps articulatory features of speech sounds. *Proc. Natl. Acad. Sci. USA*, 103(20): 7865-70, 2006.
- [9] S.M. Wilson & M. Iacoboni. Neural responses to non-native phonemes varying in producibility: evidence for the sensorimotor nature of speech perception. *NeuroImage*, 33(1): 316-25, 2006.
- [10] L. Fadiga, L. Craighero, G. Buccino & G. Rizzolatti. Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience*, 15: 399-402, 2002.
- [11] I.G. Meister, S.M. Wilson, C. Deblieck, A.D. Wu & M. Iacoboni. The Essential Role of Premotor Cortex in Speech Perception. *Current Biology*, 17(19): 1692-1696, 2007.
- [12] L. Romero, V. Walsh & C. Papagno. The neural correlates of phonological short-term memory: a repetitive transcranial magnetic stimulation study. *Journal of Cognitive Neuroscience*, 18(7): 1147-1155, 2006.
- [13] M. Sato, P. Tremblay & V.L. Gracco. A mediating role of the premotor cortex in speech segmentation. *Journal of Cognitive Neuroscience*, (submitted).
- [14] E. Kohler, C. Keysers, M.A. Umiltà, L. Fogassi, V. Gallese & G. Rizzolatti. Hearing sounds, understanding actions: Action representation in mirror neurons. *Science*, 297: 846-848, 2002.
- [15] J. H. Abbs & V.L. Gracco. Control of complex motor gestures: orofacial muscle responses to load perturbations of lip during speech. *Journal of Neurophysiology* 51: 705-23, 1984.
- [16] J.W. Folkins & J.H. Abbs. Lip and jaw motor control during speech: responses to resistive loading of the jaw. *Journal of Speech and Hearing Research*, 18(1): 207-219, 1975.
- [17] V.L. Gracco & J.H. Abbs. Dynamic control of perioral system during speech: Kinematic analyses of autogenic and nonautogenic sensorimotor processes. *Journal of Neuroscience*, 54, 418-432, 1985.
- [18] S. Shaiman, & V.L. Gracco. Task-specific sensorimotor interactions in speech production.

Experimental Brain Research, 146: 411-418 2002.

[19] D.W. Purcell & K.G. Munhall. Compensation following real-time manipulation of formants in isolated vowels, *Journal of the Acoustical Society of America*, 119: 2288-2297, 2006.

[20] S. Tremblay, D.M. Shiller & D.J. Ostry. Somatosensory basis of speech production. *Nature*, 423: 866-869, 2003.

[21] J.F. Houde & M.I. Jordan. Sensorimotor adaptation of speech I: Compensation and adaptation. *Journal of Speech, Language & Hearing Research*, 45: 295-310 2002.

[22] D.M. Shiller M. Sato, V.L. Gracco & S.R. Baum, S.R. Perceptual recalibration of speech sounds following speech motor learning. *Journal of the Acoustical Society of America*, in press.

[23] S.M. Nasir & D.J. Ostry. Speech motor learning in profoundly deaf adults. *Nature Neuroscience* 11: 1217 – 1222, 2008.

[24] S. Brown, R.J. Ingham, J.C. Ingham, A.R. Laird & P.T. Fox. Stuttered and fluent speech production: an ALE meta-analysis of functional neuroimaging studies. *Human Brain Mapping*, 25(1):105-117, 2005

[25] P. Indefrey & W.J.M. Levelt. The spatial and temporal signatures of word production components. *Cognition*, 92:101-144, 2004.

[26] V.L. Gracco, P. Tremblay & G.B. Pike. Imaging speech production using fMRI. *Neuroimage*, 26: 294-301, 2005.

[27] G. Hickok & D. Poeppel, D. The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8: 393-402, 2007.

[28] M.H. Fischer & R.A. Zwann, R. A. Embodied language: A review of the role of the motor system in language comprehension. *The Quarterly Journal of Experimental Psychology*, 61(6): 825-850, 2008.