

Speech Perception Does Not Rely on Motor Cortex

Response to D'Ausilio et al.

Gregory Hickok

Center for Cognitive Neuroscience, Department of Cognitive Sciences, University of California, Irvine

greg.hickok@uci.edu

D'Ausilio and colleagues [1] report a fascinating new study showing that stimulation of human motor cortex (via TMS) directly affects the perception of speech sounds. TMS was applied to the lip or tongue areas of M1 while participants were asked to identify speech sounds that either involved prominent lip articulation, [b] and [p], or prominent tongue articulation, [d] or [t]. They found a double-dissociation: relative to a non-stimulation baseline, participants were faster to indicate that they heard a lip-related sound when TMS was applied to motor lip areas, and faster to indicate that they heard a tongue-related sound when TMS was applied to motor tongue areas. The authors conclude that "motor structures provide a specific functional contribution to the perception of speech sounds" and go on to propose "a modified 'motor theory of speech perception' according to which speech comprehension is grounded in motor circuits..." (p. 1). This latter point is critical as there is no argument over whether sensory and motor speech systems interact [2, 3]. The debate is whether motor systems are *required* for the perception of speech sounds (phonemes). The motor theory of speech perception (MTSP) holds that they are: "the objects of speech perception are the intended phonetic gestures of the speaker, represented in the brain as invariant motor commands" [4] (p. 2). This theoretical position shares much in common with mirror neuron based theories of action understanding – "... we understand action because the motor representation of that action is activated in our brain" [5] (p. 661). In these models, speech perception is often included as a form of action perception [6].

Much evidence supports the view that motor-related processes can play a role in at least some speech perception tasks [7]; the D'Ausilio et al. study adds to this literature. However, to show that motor-related processes can influence speech perception is not the same as showing that speech perception is fundamentally grounded in motor mechanisms. For example, it is well known that the perception of phonemes is influenced by word context. A speech sound that is acoustically halfway between [d] and [t] will be heard more often as [d] when paired with the vowel-consonant ending, *og*; listeners' perception is biased toward the meaningful *dog* and away from the meaningless *tog*. But this does not mean that the perception of phonemes is critically dependent on word context. After all, we are perfectly capable of perceiving *tog* when it is presented in its acoustically unambiguous form. Rather, such a result just means that lexical systems can provide top-down *modulation* of phonemic perception, or alternatively, can affect high-level decision processes [8]. It is worth noting that D'Ausilio et al. also introduced some acoustic ambiguity in their stimuli by adding noise such that baseline performance was at 75% accuracy.

In the case of speech perception, there is strong evidence that motor-related systems are not *fundamental* to speech perception, but instead, simply *modulate* the process in some way. We turn to that evidence now.

If "the objects of speech perception are the intended phonetic gestures of the speaker, represented in the brain as invariant motor commands..." [4] (p. 2) then it follows that if the motor speech system were damaged such that a patient could no longer produce speech gestures, such a patient should no longer be able to perceive speech. However, this prediction has been falsified by decades of research. First, while focal damage restricted to the left face-M1 area can produce acute severe speech production deficits (e.g., output limited to grunts), it can leave speech comprehension "completely normal" [9] (p. 443). Second, severe chronic Broca's aphasia caused by large lesions involving Broca's region, lower motor cortex and surrounding tissue can leave the patient with no speech output (e.g., case 7 in [10]) or with only stereotyped

output (e.g., cases 7, 11, 16, 17 in [10]), but with relatively preserved comprehension (~80% accuracy in matching a spoken word to a picture from a multi-item array in these five example cases [10]; note this likely includes any comprehension errors that arise from semantic deficits, thus overestimating the severity of the speech perception deficit). Third, even acute deactivation of the entire left hemisphere in patients undergoing sodium amobarbital (Wada) procedures, which produces complete speech arrest, leaves speech sound perception relatively intact (phonemic error rate < 10%) [11]. This pattern holds even when fine phonetic discrimination is required for successful comprehension (e.g., comprehending *bear* vs. *pear* which only differ by one feature) [11]. Fourth, bilateral lesions to the cortex of the anterior operculum and foot of the precentral gyrus can cause anarthria, that is loss of voluntary muscle control of speech, yet these lesions not cause speech recognition deficits [12]. Fifth, bilateral lesions to Broca's area, argued to be the core of the human mirror system [13], do not cause word level speech recognition deficits (see case 3 in [14]). Sixth, the failure of a child to develop motor speech ability, either as a result of a congenital anarthria [15] or an acquired anarthria secondary to bilateral anterior operculum lesions [16] do not cause preclude the development of normal receptive speech [15, 16]. Seventh, babies develop sophisticated speech perception abilities including the capacity to make fine distinctions and perceive speech categorically as early as 1-month of age, well before they develop the ability to produce speech [17]. Finally, species without the capacity to develop speech (e.g., chinchillas) can nonetheless be trained to perceive subtle speech-sound discriminations in a manner characteristic of human listeners, that is, categorically [18]. In short, disruption of number of levels of the motor-speech system, including its complete failure to develop, does not preclude the ability to make subtle speech sound discriminations in perception.

The facts outlined above demonstrate that the MTSP is incorrect. An alternative to the MTSP is that speech perception relies critically on auditory-related cortices in the superior temporal lobe (bilaterally) [3]. Consistent with this is the observation that perception of speech sounds results in robust and consistent bilateral physiological responses in the superior temporal gyrus, that manipulation of phonological variables modulates activity in the superior temporal sulcus bilaterally, and that bilateral damage to the superior temporal lobe produces dramatic failures of speech sound perception and auditory comprehension (in the absence of cortical deafness) [3].

If the MTSP is incorrect, how does motor information influence speech perception, as in the D'Ausilio et al. study? There are three possibilities. One is that stimulation of M1 may generate motor-to-sensory feedback, perhaps in the form of forward models, that primes the activation of certain perceptual phonemic categories. Another is that motor and perceptual information may converge on higher-order executive processes where this information is used to color decision-making processes. Both of these mechanism have been suggested to hold of top-down word-level effects of speech sound perception [8]. A third is that the observed effects are task-specific. Attention to sublexical features of speech implicates frontal motor-related systems in a way that listening to speech for comprehension does not. In fact, identifying or discriminating nonword syllables double-dissociates, both neurally and behaviorally, from the ability to use phonemic information for speech comprehension [3]. This indicates that explicit attention to sublexical speech units engages some set of processes that are not necessarily involved in normal speech processing. Perhaps motor stimulation modulates these meta-speech processes rather than perceptual mechanisms per se.

Finally, it is worth noting that the tight connection between sensory and motor speech systems can be captured by an alternative to the MTSP, namely the proposal that sensory systems are critically involved in guiding motor speech gestures—a sensory theory of speech production [3, 21]. There is strong evidence indicating that sensory-to-motor influence is critical in speech development, in maintaining normal speech articulation, and in accurately selecting and sequencing speech segments [3]. Motor-to-sensory feedback (in the form of forward models) appears to be an important feature of this circuit [21], which may be used not only to fine tune motor output but also to modulate perceptual processes. Thus, a sensory-centered speech perception system can explain the tight association between sensory and motor systems, the modulatory influence of motor activity on perception, and in addition, can explain why catastrophic destruction or developmental failure of the motor speech system does not result in catastrophic loss in the ability to understand speech.

This work was supported by NIH grants DC03681 and DC009659.

References

1. D'Ausilio, A., Pulvermuller, F., Salmas, P., Bufalari, I., Begliomini, C., and Fadiga, L. (2009). The motor somatotopy of speech perception. *Curr Biol*.
 2. Hickok, G., Buchsbaum, B., Humphries, C., and Muftuler, T. (2003). Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. *Journal of Cognitive Neuroscience* 15, 673-682.
 3. Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat Rev Neurosci* 8, 393-402.
 4. Liberman, A.M., and Mattingly, I.G. (1985). The motor theory of speech perception revised. *Cognition* 21, 1-36.
 5. Rizzolatti, G., Fogassi, L., and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat Rev Neurosci* 2, 661-670.
 6. Rizzolatti, G., and Arbib, M. (1998). Language within our grasp. *Trends in Neurosciences* 21, 188-194.
 7. Galantucci, B., Fowler, C.A., and Turvey, M.T. (2006). The motor theory of speech perception reviewed. *Psychon Bull Rev* 13, 361-377.
 8. Myers, E.B., and Blumstein, S.E. (2008). The neural bases of the lexical effect: an fMRI investigation. *Cereb Cortex* 18, 278-288.
 9. Terao, Y., Ugawa, Y., Yamamoto, T., Sakurai, Y., Masumoto, T., Abe, O., Masutani, Y., Aoki, S., and Tsuji, S. (2007). Primary face motor area as the motor representation of articulation. *J Neurol* 254, 442-447.
 10. Naeser, M.A., Palumbo, C.L., Helm-Estabrooks, N., Stiassny-Eder, D., and Albert, M.L. (1989). Severe nonfluency in aphasia: Role of the medial subcallosal fasciculus and other white matter pathways in recovery of spontaneous speech. *Brain* 112, 1-38.
 11. Hickok, G., Okada, K., Barr, W., Pa, J., Rogalsky, C., Donnelly, K., Barde, L., and Grant, A. (2008). Bilateral capacity for speech sound processing in auditory comprehension: evidence from Wada procedures. *Brain Lang* 107, 179-184.
 12. Weller, M. (1993). Anterior opercular cortex lesions cause dissociated lower cranial nerve palsies and anarthria but no aphasia: Foix-Chavany-Marie syndrome and "automatic voluntary dissociation" revisited. *J Neurol* 240, 199-208.
 13. Craighero, L., Metta, G., Sandini, G., and Fadiga, L. (2007). The mirror-neurons system: data and models. *Prog Brain Res* 164, 39-59.
 14. Levine, D.N., and Mohr, J.P. (1979). Language after bilateral cerebral infarctions: role of the minor hemisphere in speech. *Neurology* 29, 927-938.
 15. Lenneberg, E.H. (1962). Understanding language without ability to speak: a case report. *Journal of Abnormal and Social Psychology* 65, 419-425.
 16. Christen, H.J., Hanefeld, F., Kruse, E., Imhauser, S., Ernst, J.P., and Finkenstaedt, M. (2000). Foix-Chavany-Marie (anterior operculum) syndrome in childhood: a reappraisal of Worster-Drought syndrome. *Dev Med Child Neurol* 42, 122-132.
 17. Eimas, P.D., Siqueland, E.R., Jusczyk, P., and Vigorito, J. (1971). Speech perception in infants. *Science* 171, 303-306.
 18. Kuhl, P.K., and Miller, J.D. (1975). Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science* 190, 69-72.
 19. Hickok, G. (in press). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*.
 20. Lotto, A.J., Hickok, G.S., and Holt, L.L. (in press). Reflections on Mirror Neurons and Speech Perception. *Trends Cogn Sci*.
 21. Guenther, F.H., Hampson, M., and Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review* 105, 611-633.
-

Speech Perception May Causally Depend from the Activity of Motor Centers

Reply to Hickok

Alessandro D'Ausilio,¹ Friedemann Pulvermüller,² Paola Salmas,³ Ilaria Bufalari,¹ Chiara Begliomini,¹ and Luciano Fadiga^{1,3}

¹DSBTA - Section of Human Physiology, University of Ferrara, Ferrara, Italy

²Cognition and Brain Sciences Unit, Medical Research Council, Cambridge, UK

³IIT, The Italian Institute of Technology, Genova - Italy

We thank our colleague Gregory Hickok for his comment on our paper. Discussion is always a good thing and plurality of opinion is usually, too. However, Dr. Hickok thinks "we are perfectly capable of perceiving *tog* when it is presented in its acoustically unambiguous form." In our study we show exactly that, for this to happen, activity in the motor system needs to be consistent with the information reaching the temporal lobes. If this condition is not met, we may mistake *tog* for *tod*. As this happens, in spite of the obvious fact that the ears are indeed not attached to the motor system, we concluded that motor systems interact with superior-temporal cortex in the speech perception process. One may conceptualize the underlying mechanisms as similar to attentional influences, stemming from the bidirectional feedback and feedforward connections [1] between superior-temporal and motor systems, and leading to an enhancement of superior-temporal activation as a consequence of the joint system they encompass [2].

Another important point we would like to stress is that, although we apply TMS on M1, we explicitly state in our paper that areas adjacent to M1 may be critically involved in speech perception. For example, premotor cortices could be involved in this effect. The striking finding is, however, that the facilitation and disfacilitation is manifest in a somatotopic manner, yielding double dissociations on accuracies and reaction times, thus demonstrating a causal relationship between motor and acoustic mechanisms. This causal relationship provides a physiological explanation to the previous TMS study of our group showing that speech listening induces a specific facilitation on motor cortex [3].

In sharp contrast with this view, old neurological models [4, but see 5 for a critical historical commentary], and equally the proposal by Hickok [6], have denied a necessary role of the motor system in speech perception. This is in contrast with evidence from the aphasia literature, where it had been known for a long time that aphasia, even if its underlying lesion is restricted to the frontal cortex, is a general multimodal deficit affecting both the production of speech and its perception and comprehension [7]. Clinical tests for selecting aphasics from other brain-damaged individuals include, thus, speech comprehension test [8]. Furthermore, aphasic patients generally exhibit abnormalities in speech perception [9], especially a deficit in phoneme identification, in tasks such as the one used in our study [10]. We should also stress that the hypothesis of perceptual relevance of motor systems requires precise experiments addressing this issue. However, Dr. Hickok refers to negative evidence that did not explicitly test perceptual relevance of motor centers, but rather are based on anecdotic reports or clinical tests at best [see ref. 11 for an interesting demonstration of why Broca's aphasics usually show intact comprehension in standard clinical tests, although they are impaired in such ability]. For example, what Dr. Hickok considers an index of preserved comprehension (80% of accuracy) is, in our view, a really relevant deficit. Finally, although we consider patients studies as strongly informative on brain function, we should keep in mind the fact that it is often extremely difficult to generalize these data to situations not specifically tested by a given study.

Dr. Hickok proposes three alternative interpretations to explain our data, that we summarize as follows: 1. motor to sensory flow (activation of forward models); 2. existence of a "third" decision area gathering information from sensory and motor cortices; 3. TMS targeted attentional processes towards phonological features. The first explanation is actually our interpretation. And we clearly stated such position in the

following sentences:

"It should be stressed, however, that our finding does not prove that M1 is directly involved in speech perception."

"We are not suggesting, however, that the motor cortex is an area for phonological discrimination per se; rather, we favor the idea that it might be part of a larger network."

"We propose that TMS of M1 might have unbalanced the network dynamics of action-perception circuits, likely involving motor, premotor, and temporo-parietal areas" [10, pp. 383].

The second interpretation reminds 18-19th century models of the human mind, in the sense that requires an additional functional module that is not clear how it even relates to Hickok's own model. However, if this is to be taken as a worthwhile working hypothesis it needs better definition and novel experiments—otherwise we don't see any substance in this proposal. The third, concerning the possible task-specificity of motor circuits recruitment in speech perception is ruled out by the very rapid spreading of neural activity to the motor system in passive speech perception [12, 13]. Importantly, this activation of motor systems occurs also in the absence of an attention-demanding task.

One may still want to claim, "The temporal lobe perceives speech while the motor system only helps." However, we think that this position stems from old-fashioned philosophies about the nature of brain areas as a modular input or output processors. As we point out in our paper, advances in the brain sciences in the last twenty years have taught us that neuronal assemblies encompass motor and perceptual "modules" of the brain and build distributed functional systems to which especially the motor system makes an eminent contribution [14]. Our study proves that, in the language domain, this general insight is as relevant and as valid as elsewhere. To this respect we would like to stress the fact that the real innovation provided by our study is represented by the double dissociation associated to the different phonemes (or syllables). In other words, what we demonstrate here is that the motor system does not exert a merely unspecific facilitation on perceptual areas, but that this facilitation is specific and is organized according to motor somatotopy. Thus, specific motor-perceptual channels seem to exist in the brain and these channels work by associating the acoustic property of, e.g., the speech sound /b/ with the motor representation of the articulatory gesture leading to the production of the same speech sound in the listener's motor brain. We see this finding very close to the Liberman's idea of motor perception and we felt ourselves obliged to recognize the intellectual merit of his intuition.

Distributed systems with a strongly linked action and perception subcomponents explain patterns of deficits in aphasia, especially dissociations between motor and perceptual impairments in case of lesion of the distributed neuronal assemblies at their acoustic or motor ends [15, 16]. Ultimately, as a distributed circuit needs to receive sensory input and control motor output, cutting of these afferent and efferent connections does explain the occasionally observed unimodal deficits mentioned in Hickok's contribution. By no means do these dissociations prove the modular nature of the language system. Lesion evidence argues in favour of a distributed systems account [17]. In sum, we do not think that Hickok's proposal provide reasonable arguments for rejecting functional interactions between motor and language systems, speech perception systems included. Nor it is justified to make the drastic claim expressed by the title of his Correspondence, in light of our and others' recent findings. Hickok's point of view is a point of view and cannot gain evidence through his apodictic title.

References

1. Saur, D., Kreher, B. W., Schnell, S., Kummerer, D., Kellmeyer, P., Vry, M. S., Umarova, R., Musso, M., Glauche, V., Abel, S., Huber, W., Rijntjes, M., Hennig, J., & Weiller, C. (2008). Ventral and dorsal pathways for language. *Proc Natl Acad Sci U S A*, 105, 18035-18040.
2. Uppenkamp, S., Johnsrude, I. S., Norris, D., Marslen-Wilson, W., & Patterson, R. D. (2006). Locating the initial stages of speech-sound processing in human temporal cortex. *Neuroimage*, 31, 1284-1296.
3. 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.
4. Lichtheim, L. (1885). Über Aphasie. *Deutsches Archiv für Klinische Medizin* 36, 204-268.
5. Berker, E., Berker, A., & Smith, A. (1986). Translation of Broca's 1865 report. Localization of speech in the third left frontal convolution. *Arch Neurol* 43, 1065-1072.
6. Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat Rev Neurosci*,

8, 393-402.

7. Rosenbek, J. C., LaPointe, L. L., & Wertz, R. (1995). *Aphasia: A clinical approach* (2nd Edition ed.). Boston: College-Hill Press.
 8. De Renzi, E., and Vignolo, L. (1962). The Token Test: a sensitive test to detect receptive disturbances in aphasics. *Brain*, 85, 665-678.
 9. Basso, A., Casati, G., and Vignolo, L.A. (1977). Phonemic identification defect in aphasia. *Cortex*, 13, 85-95.
 10. D'Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., and Fadiga, L. (2009). The Motor Somatotopy of Speech Perception. *Curr Biol*, 19, 381–385.
 11. Moineau, S., Dronkers, N. F., Bates, E. (2005). Exploring the processing continuum of single-word comprehension in aphasia. *J Speech Lang Hear Res*, 48, 884-96.
 12. Pulvermüller, F., Shtyrov, Y., and Ilmoniemi, R.J. (2003). Spatio-temporal patterns of neural language processing: an MEG study using Minimum-Norm Current Estimates. *Neuroimage*, 20, 1020-1025.
 13. Pulvermüller, F., and Shtyrov, Y. (2009). Spatio-temporal signatures of large-scale synfire chains for speech: MEG evidence. *Cereb Cortex*, 19, 79-88.
 14. Rizzolatti, G., Fogassi, L., and Gallese, V. (2006). Mirrors of the mind. *Sci Am*, 295, 54-61.
 15. Pulvermüller, F., and Preissl, H. (1991). A cell assembly model of language. *Network*, 2, 455-468.
 16. Garagnani, M., Wenekers, T., and Pulvermüller, F. (2008). A neuroanatomically-grounded Hebbian learning model of attention-language interactions in the human brain. *Eur J Neurosci*, 27, 492-513.
 17. Pulvermüller, F. (1999). Words in the brain's language. *Behav Brain Sci*, 22, 253-336.
-