

Proposed Topic: Neural Bases for the Motor Theory of Speech Perception

A theory of speech perception that maps between auditory input and articulatory features was most famously proposed by Liberman et al. (1967). Their theory was later revised by the same authors (Liberman & Mattingly, 1985) to account for new discoveries in comparative evolutionary biology and developmental psychology. Although not the first to propose that gestural features were the fundamental unit of the speech code, their “motor theory of speech perception” has become one of the most cited (though perhaps predominately as a counter-example to an auditory explanation). Other theories are very similar in their approach, though they go by different names, such as Carol Fowler’s “direct realism” (see Diehl, Lotto, & Holt, 2004, for a review of various theories of speech perception). The “motor theory of speech perception” (or, hereafter, sensorimotor theory) makes two specific and empirically testable predictions about the nature of spoken language processing. First, the fundamental units of spoken language are shared between speech production and perception; and, second, these underlying units consist of the articulatory features involved in producing speech.

Why is your topic interesting? What is its significance to speech and hearing?

Such a sensorimotor theory is attractive for at least two reasons, which appear largely unaddressed in other, auditory-only theories: First, it explains how highly variable acoustic signals give rise to invariant phonemic percepts. Second, it provides a parsimonious way to link the elements of spoken language production and perception into a single system. Sensorimotor theories of speech perception are an extremely interesting and open avenue of research. There has not been a great deal of empirical work aimed at testing their specific predictions, leaving room

for a considerable amount of future innovation. This is especially curious considering how often sensorimotor theories are invoked as an explanation (or possible counter explanation to satisfy a sense of “appealing to both sides”) for speech perception phenomena, and in fact language and linguistic evolutionary theories in general. The practical implications from a rigorous test of the motor theory are substantial. A better understanding of how speech production and speech perception systems interact would be very beneficial for formulating more specific hypotheses in understanding language disorders such as stuttering, dyslexia, apraxia and aphasia. Similarly, a better understanding of not only the interaction between perception and production, but knowledge of how spoken language was represented fundamentally, could potentially facilitate clinicians treatment programs of these same disorders. Not surprisingly, understanding the neural representation of the fundamental units of (spoken) language would be of substantial didactic interest to linguists and cognitive scientists, who could develop more accurate and predictive models of language processing. Knowing whether there was any concomitant benefit between production and perception training, pedagogues of foreign language acquisition (or, in fact, first language acquisition in the cases of language-learning impairment and specific language impairment) could know the relative benefit of emphasizing training in articulation or comprehension.

Moreover, additional perspective on how integrated sensorimotor networks can carry out complicated social behaviors such as spoken language provides a valuable corollary for psychologists and neuroscientists studying the brain basis of imitative learning, including what role may be played by the putative mirror-neuron system in speech (e.g. Rizzolatti & Arbib, 1998). Additionally, the congruence between distal events and proximal sensory representations in an integrated sensorimotor network offers insights into how the brain might be wired for other

perceptual tasks, such as the recognition of conspecifics or making judgments about the emotions and intentions of others, linking the study of speech to social and affective neuroscience and research on the theory of mind. A more comprehensive description of linguistic representation could likewise offer further insight into the neural changes that allowed for such sensorimotor integration of information, ultimately giving rise to language during human evolution, and will thus be broadly interesting to anthropologists and biologists. Ultimately, investigating sensorimotor theories of speech perception is not a question of how speech works in isolation, or how the brain works to support a specific task, but rather supports broader theories about brain evolution to support complex integration between sensory, cognitive, and behavioral phenomena.

How does your topic relate to others we have covered or will cover in class?

Implications of the sensorimotor theory of speech perception have been brought up (or could have been brought up) in every unit we have considered so far. First, in studying the dorsal cochlear nucleus, we considered how sensorimotor and auditory information were actively integrated, even at relatively peripheral levels, for the purposes of sound localization. Just as cats and other animals consider the location of their pinna with respect to the location of putative sounds, it seems likely that humans could (even minimally) use knowledge about articulation, whether innate or learned, to help decompose speech sounds into specific phonological percepts. Second we considered quantitative methods for neural coding. Although not explicitly addressed during this unit, I personally believe some of the methodologies we discussed are critical for determining the extent of sensorimotor integration in speech perception. Approaches such as mutual information could be used to compare between patterns of cortical activation in auditory and motor regions to the original speech stimulus, indicating how these regions were coding for speech sounds and whether the motor or auditory representations were more “invariant.” The

question of the sensorimotor theories came up frequently in the sections on speech motor control, and although sensorimotor theories were not specifically implicated in speech *perception* based on what Dr. Guenther presented to us, there is no reason to believe that the integrated functional circuit he described between auditory and motor regions for speech *production* could not similarly underlie perceptual tasks. The relevance to the cortical representation of language is self-evident; needless to say it is especially interesting to consider what are the necessary and sufficient qualities of sensorimotor theories, especially based on behavioral changes resultant from brain damage. Similarly, left-right differences in auditory cortex may be particularly relevant to the sensorimotor theory: Are there different patterns of connectivity between superior temporal regions and (pre)motor cortices? How do these differences correlate with known auditory cortex asymmetries, and could they thus be also involved in any overall lateralization of language function? Sensorimotor theories of language finally bear on the study of auditory-processing disorders of language. By learning more about the fundamental representation of speech sounds, we would be better poised to test whether these disorders result from specific perception, integration, or computation deficits.

What is already known about the topic? What are some of the key issues in this research area?

The neural bases of a sensorimotor theory of speech perception have only begun to be developed in the last decade or so. Electrophysiological recordings and neuroimaging in monkeys have revealed activation of frontal units during both the perception and production of sounds and communicative actions (e.g. Gil-da-Costa et al, 2006; Ferrari et al, 2003; Kohler et al, 2002). Some have proposed this putative mirror-neuron system may underlie language comprehension and/or acquisition (Rizzolatti & Arbib, 1998). Additionally, some behavioral work has been done in humans linking perception abilities to speech production, including

priming (Roelofs, Özdemir, & Levelt, 2007) and habituation (Cooper, 1979). Other evidence from perceptual experiments designed to distinguish spectral (auditory) processing with gestural (motor) processing (Fowler, 2006). Brain imaging techniques have recently begun to show considerable evidence of overlap in motor region activity between speech perception and production tasks (outlined below).

Are there any recent technical developments that foster new approaches to the topic?

Recent advances in noninvasive neuroimaging have revitalized interest in sensorimotor theories of speech perception in particular and unifying theories of action-perception in general. Recall, the specific predictions of a sensorimotor theory are that (1) underlying representation is shared between perception and production systems, and (2) the nature of this representation is articulatory. Wilson et al. (2004) used fMRI to localize cortical regions involved in the production of speech syllables. The subsequently presented the same subjects with the same syllables presented auditorily and discovered a significant overlap in motor region activity during both perception and production. Pulvermüller et al. (2006) substantially advanced this research by mapping cortical response to the production of specific phonemes (/p/, produced with the lips, and /t/, produced with the tongue). They, too, imaged the same subjects during a speech perception task to discover that the lip region was significantly more activated during the perception of /p/, whereas the tongue region was significantly more activated during the perception of /t/. Having established that the same motor region is implicated in both the perception and production of speech, other authors questioned how this region might be involved. Using TMS, several groups have demonstrated that premotor stimulation impairs speech perception more than nonspeech auditory perception (Meister et al. 2007), and that different regions of the IFG are responsible for phonological or semantic processing (Gough,

Nobre, & Devlin, 2005). Combining PET and TMS, Watkins & Paus (2004) showed that speech perception increases the excitability of the motor system involved in speech production.

Together, these studies have suggested there might be something substantial to the perception-production link that is intimately coupled with perceptual abilities. However, it remains to be shown in future studies whether the major premises of the motor theory are confirmed: the same information is being encoded in motor cortices during speech perception and production tasks, and whether the nature of this representation is always articulatory.

Selected (Annotated) Bibliography

Background Reading:

Galantucci, B., Fowler, C.A., & Turvey, M.T. (2006) "The motor theory of speech perception reviewed" *Psychonomic Bulletin and Review*, 13, 361-377.
(General discussion of the changes and evidence for the motor theory since 1968 to now).

Proposed Discussion Articles:

Wilson, S.M., Saygin, A.P., Sereno, M.I., & Iacobini, M. (2004) *Nature Neurosci.*, 7, 701-702.
(fMRI study of motor cortex activation in both production and perception of speech.)

Pulvermüller, F., Huss, M., Kheriff, F., Moscoro del Prado Martin, F., Hauk, O. & Shtyrov, Y. (2006) "Motor cortex maps articulatory features of speech sounds." *Proceedings of the National Academy of Sciences*, 103, 7865-7870.
(fMRI studying showing lip cortex activity for perception of /p/ and tongue for /t/)

Watkins, K. & Paus, T. (2004) "Modulation of motor excitability during speech perception: The role of Broca's area." *Journal of Cognitive Neuroscience*, 16, 978-987.
(PET and TMS to localize speech perception motor excitability).

Meister, I.G., Wilson, S.M., Deblieck, C., Wu, A.D., & Iacobini, M. (2007) "The essential role of premotor cortex in speech perception." *Current Biology*, 17, 1692-1696.
(Impaired speech but not nonspeech perception during TMS stimulation of PFC)

Additional Articles of Interest:

Rizzolatti, G & Arbib, M.A. (1998) "Language within our grasp." *Trends in Neuroscience*, 5, 188-194.

(Discusses role of mirror neuron system in perception-action links for speech processing).

Fowler, C.A. (2006) "Compensation for coarticulation reflects gesture perception, not spectral contrast." *Perception & Psychophysics*, 68, 161-177.

(Behavioral study suggesting gestures are the units of encoding, not auditory info only).

Fowler, C.A. (1996) "Listeners do hear sounds, not tongues." *Journal of the Acoustical Society of America*, 99, 1730-1741.

(Response to evidence suggested by critics to disconfirm the sensorimotor theories)