What happens to the motor theory of perception when the motor system is damaged?

Abstract: Motor theories of perception posit that motor information is necessary for successful recognition of actions. Perhaps the most well known of this class of proposals is the motor theory of speech perception, which argues that speech recognition is fundamentally a process of identifying the articulatory gestures (i.e., motor representations) that were used to produce the speech signal. Here we review neuropsychological evidence from patients with damage to the motor system, in the context of motor theories of perception applied to both manual actions and speech. Motor theories of perception predict that patients with motor impairments will have impairments for action recognition. Contrary to that prediction, the available neuropsychological evidence indicates that recognition can be spared despite profound impairments to production. These data falsify strong forms of the motor theory of perception, and frame new questions about the dynamical interactions that govern how information is exchanged between input and output systems.

Keywords: speech perception, language comprehension, mirror neuron system, apraxia, motor theory of perception

1 Introduction

An important insight reflected in Michael Arbib’s monograph How the Brain got Language is that recognition of perceptual information is not a passive pro-
cess and does not occur simply for its own sake, but ultimately in the service of behavior. The recent and widespread interest in how cognitive and perceptual processes are inherently active processes has been an important force shaping new approaches for understanding how the human mind works. This commentary focuses on an issue that runs largely in parallel to Arbib’s theory: If recognition can be understood as an active process whereby implicit ‘hypotheses’ are generated about the nature of the percept, then what is the format of the information over which those ‘hypotheses’ are formulated? Is the format of the representations that are involved in recognition local to the perceptual process itself, or are qualitatively different representations involved in the recognition process?

Motor theories of perception, which have recently been tied to the Mirror Neuron Hypothesis, propose that the format of at least some of the information over which such ‘hypotheses’ are formulated is motor. Rizzolatti and Arbib (1998) adopted the basic components of the motor theory of speech perception in the context of mirror neurons and a theory about the evolutionary history of language. However, the motor theory of perception does not figure in Arbib’s 2012 book, How the Brain got Language, and his core argument about the evolution of language, which we do not address, could be true even if strong forms of the motor theory of speech perception are false. Arbib (2012) does discuss the notion of “language parity,” which is the idea that, to quote Arbib (p. 37), “… the meaning understood by the receiver is (at least approximately) the meaning intended by the sender” – but Arbib’s theory is not committed to the view that language parity is achieved over motor information, and his new formulation of the evolutionary history of language does not involve the claim that the retrieval of motor activation is a necessary and intermediary step in action recognition:

“I would suggest that the understanding of all actions involves general mechanisms that need not involve the mirror system strongly – but that for actions that are in the observer’s repertoire, these general mechanisms may be complemented by activity in the mirror system that enriches that understanding by access to a network of associations linked to the observer’s own performance of such actions (p. 142).”

The focus of our article is not on mirror neurons per se, but on the motor theory of action recognition, and specifically, the motor theory of speech perception (for critical discussion of mirror neurons in humans see Dinstein et al. 2008; Hickok 2009). Since so much of recent work on mirror neurons has been aligned with motor theories of perception, we take this opportunity, afforded by the publication of Arbib’s monograph, to offer a re-evaluation of motor theories of perception.

Our goal is two-fold. The first, and more minor goal is to argue that the inference from the available evidence to the motor theory of perception is premature.
because we lack a theory of the dynamics of activation spread between input and output systems. Only in the context of explicit assumptions about how activation spreads between sensory and motor processes, and how that spreading of activation may be ‘gated’ (if at all), can inferences be drawn about the format of the information that is causally involved in recognition.

Our second, and principal goal, is to review the available neuropsychological evidence as it relates to the question formulated in our title: What happens to the motor theory of perception when the motor system is damaged? The critical test of whether motor processes play a constitutive (i.e. necessary and intermediary) role in recognition processes is provided by patient evidence. As we will review, motor processes can be impaired (at multiple levels) while recognition processes remain intact (again, at multiple levels). This empirical fact sharpens the discussion of what kind of dynamical assumptions are implied by the evidence that has been taken as the primary support for motor theories of perception in the first place.

2 Motor theories of perception – A fast survey

The motor theory of speech perception (Liberman et al. 1967; Liberman and Mattingly 1985) proposed that speech recognition is fundamentally a process of recognizing or matching the motor programs that the speaker used to generate the speech sounds. In their original theory, Liberman and colleagues (1967) argued that humans’ perception of speech is based on articulatory gestures that are shared by both the speaker and listener and that a listener perceives speech when his or her actual articulatory representations are activated by the speaker’s utterances. In a revised theory, Liberman and Mattingly (1985) maintained three claims: i) speech processing is special (i.e. unique to humans); ii) perceiving speech is perceiving the vocal gestures of the speaker (versus acoustic information); and iii) the motor system is necessary for perceiving speech. On the basis of those theoretical commitments, it was argued that speech perception and speech production are part of one motor module and share neural processes.

Originally, the Motor Theory of Speech Perception had mixed reviews and a few decades after its inception there seemed to be no direct evidence to support it (see Galantucci et al. 2006 for review), and several findings were problematic for the theory. For instance, the fact that chinchillas can discriminate human speech indicates that a production system is not necessary for recognition to occur (Kuhl and Miller 1975). So the question becomes whether humans use their production system for recognition, and if so, why? Evidence from human development indicated that infants discriminate sounds that they cannot yet produce (Eimas et al.
so it is not clear that there is any coherent developmental story to tell for why the production system would be used for recognition. However, the subsequent discovery of ‘mirror neurons’ rejuvenated interest in the motor theory of speech perception. The restorative effect that mirror neurons had on interest in motor theories of perception is illustrated in Figure 1, which plots the number of citations per year to Liberman and colleagues’ (1967) ‘Perception of the speech code,’ from 1968 to present.

Fast-forward to the 2000’s and there are a number of captivating demonstrations that the motor system is automatically engaged during perception of speech sounds. Fadiga and colleagues (2002) found that there was an increase in motor-evoked potentials (MEPs) recorded from the listener’s tongue muscles while listening to speech. In another study, using repetitive transcranial magnetic stimulation (TMS) to temporarily inactivate premotor cortex, Meister and colleagues (2007) observed that disruption of premotor cortex affected the discrimination

Fig. 1: The recovery of interest in the motor theory of speech recognition was fueled by the discovery of mirror neurons in the 1990s (e.g. di Pellegrino et al. 1992; for reviews see Rizzolatti et al. 2001; Rizzolatti and Craighero 2004). This is illustrated in the figure, which plots the number of citations per year to Liberman and colleagues (1967) ‘Perception of the speech code,’ from 1968 to present.
of stop consonants embedded in noise but did not affect color discrimination. Also using repetitive TMS, Möttönen and Watkins (2012) found that temporarily disrupting the lip representations in left motor cortex impaired subjects’ ability to discriminate between synthetic speech sounds that are lip-articulated (e.g. “ba” vs. “da”) but not sounds that are not lip-articulated (e.g. “ka” vs. “ga”), consistent with the core assumptions of the motor theory of speech perception.

2.1 Are the available data evidence for the motor theory of perception?

Many different types of data have been argued to support the existence and function of a mirror system in monkeys and humans, including neurophysiology, TMS, fMRI, EEG, lesion studies, and behavior in healthy individuals. It is beyond the scope of the current discussion to review those findings (for theory and review, see Fadiga et al. 1995; Hesslow 2002; Kiefer and Pulvermüller 2012; Rizzolatti and Craighero 2004; Scheerer 1984; and Arbib’s *How the Brain got Language*).

Here we highlight and re-assess a few example findings that form the strongest evidence for a Motor Theory of Speech Perception. Consider the TMS data reviewed above indicating that TMS to motor areas modulates MEPs recorded from the tongue while listening to speech sounds. The critical issue is whether those data distinguish between the motor theory of speech perception, and an alternative explanation according to which activation spreads from perceptual levels of processing through to motor processes. There are different ways in which such an alternative could be formulated. For instance, it could be argued that the dynamics of the system are such that activation propagates to the motor system only after the stimulus has been recognized as such. Alternatively, it could be argued that activation cascades forward to the motor system from input levels of processing prior to completion of processing at those input levels. The broader point is that a range of alternative accounts can be formulated to explain why the motor system is activated during perception.

There is a qualitative difference between the findings of Fadiga and colleagues (2002) which showed only that the motor system is activated during perception, and the findings of (for instance) Meister and colleagues (2007), who showed that inactivation of motor structures impairs perception. However, much the same (and unresolved) issue arises here regarding the dynamics of activation spread between sensory and motor systems. The argument that the data from Meister and colleagues supports the motor theory of perception carries with it the strong assumption that the effects of TMS are local to the inactivated region. In other words, the assumption must be made that inactivating premotor cortex
with TMS does not perturb processing in sensory regions that are known to be connected to premotor cortex.

2.2 Why assume a motor theory of perception in the first place?

The original reasoning was that using the production system ‘solved’ the invariance problem: there is no one-to-one mapping between acoustic information and perception. Specifically, the argument was that because gestural representations are (putatively) stable, if the acoustic information were ‘recognized’ through a match to those gestural representations, then the invariance problem would be solved. One problem with this approach is that in order to match the input to a gesture representation, the acoustic information must be parsed – which would seem to introduce some circularity into the motor theory of speech perception. That is, you need to know from acoustics alone which gestures are intended before you can combine the gestural information with the correct acoustics. Setting that consideration aside, and on closer examination, the invariance problem between acoustic information and gestures turns out to be worse than the invariance problem between acoustic information and perception. Acoustic variability is a by-product of co-articulation. Thus, the supposition that gestural representations are ‘stable’ is itself difficult to maintain. Without that assumption, the motor theory of perception loses the (putative) explanatory advantage that formed an important part of the motivation for the theory in the first place. As noted by Lotto and colleagues (2009), mirror neurons do not add anything here in solving the co-articulation problem, as they do not provide insight into the mapping between production and perception. So, bracketing for the moment the evidential status of the motor theory of perception, it is not at all clear that the theory meets its own explanatory burden (for cogent discussion of these conceptual issues see Lotto et al. 2009).

So then, what is the evidential status of motor theories of perception? As Hickok (2010) noted, the majority of studies provide only correlational evidence for the role of the motor speech system in speech perception. As those data are suggestive but not distinguishing, we can set them aside. The TMS inactivation

1 The same issue arises in considering how the motor system could come to be activated in a somatotopic manner according to linguistic meaning (e.g. ‘kick’ vs. ‘lick’; see Mahon and Caramazza 2008 for discussion). It is not clear how such motor activation could be involved in the initial parse of what that meaning is, given that its somatotopic activation indicates the meaning has already (at some level) been extracted from the input.
data discussed above are compelling but suffer from too many unknowns about how a TMS stimulus is propagated through a network. This leaves us with neuropsychological evidence. As Toni and colleagues (2008: 72) noted: “... only when it can be shown that speech recognition is no longer possible if the motor areas that subserve the neuromotor commands of the articulatory gestures are lesioned, do we have reason to believe that the motor theory of speech perception is correct.” So what is the patient evidence? We first consider evidence as it bears on a motor theory of manual action recognition, and then consider evidence in the language domain as it relates to the motor theory of speech perception.

3 Patient evidence and the motor theory of manual action recognition

Limb apraxia is a deficit in using objects accurately, which cannot be attributed to basic sensory or motor impairments such as muscle weakness or a failure to recognize the object. Findings from the study of apraxia have been used to argue both for and against motor theories of perception. For example, Pazzaglia and colleagues (2008) found that patients with buccofacial apraxia had differential impairments for recognizing mouth-compared to hand-generated sounds, while patients with limb apraxia had differential impairments for recognizing hand-generated compared to mouth-generated sounds (see also Serino and colleagues 2009 for a relevant study on action recognition). However, patients with limb apraxia who have dramatic impairments for being able to produce actions, have been observed to present with spared action recognition (Garcea et al. 2012; Negri et al. 2007; Rapsacas et al. 1995; Rumiati et al. 2001; for review, see Mahon and Caramazza 2005). Thus, an impairment for the execution of manual actions does not necessarily imply an impairment for recognition or comprehension of manual actions (see Figure 2).

4 Patient evidence and the motor theory of speech perception

There is a range of neuropsychological findings that challenge the core commitment of the motor theory of speech perception. For instance, deactivation of the entire left hemisphere through the use of sodium amobarbital (WADA test) usually results in a complete failure to produce speech. However, perception can
remain intact, as demonstrated by Hickok and colleagues (2008). In that study the phonemic discrimination error rate remained below 10% as shown by an auditory word-to-picture matching task. Thus, partial or complete disruption of the motor speech system does not necessarily result in a deficit for speech discrimination.

According to the motor theory of speech perception, lesions to the neural substrate of motor speech control, Broca's area, should result in an impairment for speech recognition. Originally, the performance of patients with 'Broca's aphasia' was argued to be evidence for the motor theory of speech perception because a subset of patients also had deficits in speech sound discrimination (e.g. Blumstein 1995). However, there are a number of reports of so-called 'Broca's aphasics' who are not impaired for speech perception (Baker et al. 1981; Blumstein et al. 1977; Moineau et al. 2005; Rogalsky et al. 2011). This empirical generalization is reinforced by the study of Hickok and colleagues (2008) of 24 individuals.
Testing the motor theory of speech perception

with non-fluent speech production subsequent to strokes that lesioned Broca’s area. In that study, damage to Broca’s area did not result in significant speech perception deficits, as tested with auditory syllable discrimination (i.e. distinguishing “ba” from “pa”) and auditory word discrimination (i.e. distinguishing “pear” from “bear”). In addition, those patients also performed well on auditory word and visual word comprehension tasks (see Figure 3, reproduced from Hickok et al. 2011, for example data).

In another relevant study, Rogalsky and colleagues (2011) studied five patients with lesions to the ‘human mirror system’ on tests of word comprehension and syllable discrimination. An inclusion criterion for patients in that study was the presence of brain damage to motor areas such as Broca’s area or the inferior parietal lobule (or both). Two subjects with fronto-parietal damage (and a spared temporal lobe) scored at ceiling on the comprehension tests and were 95% or higher on both of the discrimination tasks. The other subjects, who presented with both fronto-parietal and temporal lobe damage, were worse, but still relatively accurate, on the comprehension tasks. Rogalsky and colleagues (2011) concluded that the fact that patients with damage to the (putative) human mirror system performed well on all tests of speech comprehension and discrimination falsifies the strong form of the motor theory of speech perception. Furthermore, the fact that the worst performance on comprehension tasks was in patients with damage to the temporal lobe is consistent with the view that the temporal lobes...
are necessary and sufficient for speech perception (see Diehl et al. 2004 for review).

A final class of patients are those with Apraxia of Speech (AOS; see Darley 1968 for early cases). AOS is defined as a selective disturbance of the articulation of words resulting from a disruption to the left hemisphere motor system. While only a few cases of apraxia of speech without aphasia have been reported, the impairments that characterize AOS can occur in the absence of other language impairments. A patient with 'pure' AOS has the intent, the underlying linguistic representation, and the fundamental motor ability to produce speech, but cannot do so volitionally and may produce inconsistent, articulatory errors that are close to the target word. Critically, although patients with AOS are impaired in their ability to coordinate speech movements, their ability to perceive speech sounds (including their own) may remain unaffected (e.g. Blumstein 1991; Dronkers 1996; see also Johns and Darley 1970 and Shankweiller and Harris 1966 for early work; see Pazzaglia et al. 2008 for a counterargument in the semantic domain). For instance, Square-Storer and colleagues (1988) assessed nonspeech and auditory speech processing in a group of four patients with 'pure' apraxia of speech, who had no co-existing aphasia. Those four patients successfully completed tests of auditory comprehension, auditory retention, verbal language, and reading and writing. The assessment included seven tasks of speech recognition/discrimination in which phoneme differences were varied by manner or place of articulation, as well as voicing. In addition, internal speech discrimination tasks measured each subject’s ability to phonetically assess the structure of highly familiar monosyllabic words. The results of the study indicated that patients with ‘pure’ AOS were able to process auditory and speech stimuli as well as control subjects, whereas the aphasic patients and AOS patients with associated aphasia were impaired on the same tasks relative to controls (and relative to the ‘pure’ AOS patients). From this the authors concluded that a characteristic trait of ‘pure’ apraxia of speech is spared speech comprehension/recognition.

5 Summary and conclusion

The neuropsychological evidence that has been briefly reviewed establishes the empirical fact that speech production can be impaired (at multiple levels of analysis) while speech recognition (at multiple levels of analysis) remains intact. From this we conclude that the motor theory of speech perception is not viable, regardless of the activation evidence that has been cited in support of that hypothesis. How can this conclusion and the attendant patient data be reconciled with observations that repetitive TMS inactivating premotor cortex impairs per-
ception (e.g. Meister et al. 2007; Möttönen and Watkins 2012)? The juxtaposition of the patient evidence with those TMS findings would seem to lend support to the hypothesis that the effects of TMS are not restricted to the stimulated or deactivated region.

Where does this leave us with respect to the theory outlined by Arbib in his monograph *How the Brain got Language*? An important aspect of Arbib’s contribution is to enumerate some of the limitations of mirror neurons, and to draw a distinction between the role of motor systems in supporting analysis of the speech signal itself, from computations of the meaning of the speech signal. For instance, Arbib (2012: 281) writes:

“The resolution is to stress that the Mirror System Hypothesis views mirror neurons for words as encoding the signifier but not in general the signified. Mirror neurons for words encode the articulatory form but must be linked to other neural networks for encoding meaning.”

While Arbib’s broader position (see e.g. Arbib 2010) is largely in line with the proposal of ventral and dorsal paths for speech processing (see Hickok and Poeppel 2004, 2007), the critical issue is whether Arbib’s framework assigns a causal (i.e. necessary) role to motor systems in parsing the speech signal. Our understanding of Arbib’s new framework is that it is sufficiently circumspect on this issue so as to not be troubled by clear demonstrations of impaired production and spared recognition; this is because his theory is largely independent of motor theories of perception. In that regard, our hope would be that the types of neuropsychological evidence that we have reviewed may offer directions for further constraints within his framework, as well as within other proposals that draw on the functions of mirror neurons.

More generally, the findings we have summarized in this article indicate that the motor system is highly interconnected with, and relevant in some as yet unspecified way, to perception. That, to our knowledge, has never been denied, even by proponents of the view that speech perception occurs over (only) auditory or auditorily-relevant information. The question then becomes whether motor information can facilitate recognition of speech under degraded conditions, or provide top-down constraints that may assist in guiding the formulation of hypotheses over auditory information. In other words, ‘analysis by synthesis’ can occur in the auditory domain, and can be informed by relevant information that is represented and processed by other systems (including, but not limited to the motor system). This suggests a shift in research, from demonstrations of the ‘mere fact’ that the motor system is activated during perception, to research aimed at unpacking the processing dynamics that mediate interactions between input and output systems.
Acknowledgments: The authors would like to thank Michael Arbib, Richard Aslin, Gregory Hickok, David Kemmerer and Michael Tanenhaus for their comments on an earlier draft. Preparation of this manuscript was supported by NIH grant R21 NS076176 to BZM.

References


