Brain & Language 164 (2017) 77–105

Contents lists available at ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

Review

The hearing ear is always found close to the speaking tongue: Review of the role of the motor system in speech perception

Jeremy I. Skipper [a,*], Joseph T. Devlin [a], Daniel R. Lametti [a,b]

[a] *Experimental Psychology, University College London, United Kingdom*
[b] *Department of Experimental Psychology, University of Oxford, United Kingdom*

ARTICLE INFO

Article history:
Received 17 July 2016
Accepted 24 October 2016
Available online 5 November 2016

Keywords:
Articulation
Complex network
Phoneme
Neuroimaging meta-analysis
Speech production

ABSTRACT

Does "the motor system" play "a role" in speech perception? If so, where, how, and when? We conducted a systematic review that addresses these questions using both qualitative and quantitative methods. The qualitative review of behavioural, computational modelling, non-human animal, brain damage/disorder, electrical stimulation/recording, and neuroimaging research suggests that distributed brain regions involved in producing speech play specific, dynamic, and contextually determined roles in speech perception. The quantitative review employed region and network based neuroimaging meta-analyses and a novel text mining method to describe relative contributions of nodes in distributed brain networks. Supporting the qualitative review, results show a specific functional correspondence between regions involved in non-linguistic movement of the articulators, covertly and overtly producing speech, and the perception of both nonword and word sounds. This distributed set of cortical and subcortical speech production regions are ubiquitously active and form multiple networks whose topologies dynamically change with listening context. Results are inconsistent with motor and acoustic only models of speech perception and classical and contemporary dual-stream models of the organization of language and the brain. Instead, results are more consistent with complex network models in which multiple speech production related networks and subnetworks dynamically self-organize to constrain interpretation of indeterminant acoustic patterns as listening context requires.

Contents

* Corresponding author at: University College London, Experimental Psychology, 26 Bedford Way, London WC1H OAP, United Kingdom.
E-mail address: jskipper@lab-lab.org (J.I. Skipper).

http://dx.doi.org/10.1016/j.bandl.2016.10.004
0093-934X/

1. Introduction

Why does it matter if “the motor system” or, specifically here, brain regions supporting speech production, play a role in speech perception? It matters because, after decades of research, we still do not know how we perceive speech sounds even though this behaviour is fundamental to our ability to use language. One hindrance to this understanding has been an inability to specify how we hear sounds as particular (and putative) speech categories like phonemes or syllables. Indeed, no acoustic features have been found that can uniquely and consistently be used to characterize those units (Appelbaum, 1996; Appelbaum, 1999; Goldinger & Azuma, 2003; Port, 2010/1). Speech production systems matter in this context because it has long been proposed by theoretical models of speech perception that this problem of acoustic indeterminacy – or “lack of invariance” – can be addressed by making reference to the motor system. In particular, the motor theory of speech perception proposed that “sounds are not the true objects of perception ... rather, they only supply the information for immediate perception of the gestures” (Liberman & Mattingly, 1985). These gestures are “represented in the brain as invariant motor commands that call for movements of the articulators” and involve a “perception-production link [that] is a necessary condition for recognizing speech as speech” (Liberman & Mattingly, 1985). In contrast, the “analysis-by-synthesis” (AxS) model proposed that the motor system assists perception by providing production-based constraints on the interpretation of acoustic patterns as needed (Bever & Poeppel, 2010; Poeppel & Monahan, 2011; Skipper, Nusbaum, & Small, 2006; Stevens & Halle, 1967).

If either of these models were accurate, there are implications not only for our theories of speech perception, but also our understanding of the organization of language in the brain. That is, though neither model is neurobiologically well specified, both can be used to make inferences about the brain basis of language. The motor theory of speech perception suggests that the motor system needs to play a role in the neurobiology of speech perception. In contrast, the AxS model suggests that speech perception is more distributed in the brain with the motor system contributing dynamically in an active, constructive, or predictive manner. Neither model is consistent with “textbook” or “classical” models of the organization of language in the brain because speech production and perception are presented as separable neurobiological processes in those models, with production occurring as a result of processing in Broca’s area and comprehension as a result of processing in Wernicke’s area (a model still taught to medical students; Geschwind, 1970).

Both models also seem to be inconsistent with the most cited contemporary model of the organization of language and the brain, the “dual-stream” model of speech and language (Hickok & Poeppel, 2007; Poeppel & Hickok, 2004). Though the “dorsal stream” in this model is proposed to support “sound to action”, it does “not appear to be a critical component of the speech perception process” (Poeppel & Hickok, 2004). As support, Hickok, Costanzo, Capasso, and Miceli (2011) analysed patients with damage to Broca’s area and concluded that the “motor speech system is not necessary for speech perception” (p. 214). Other scientists seem to support this view. For example, in their review, Scott, McGettigan, and Eisner (2009) state that “the motor cortex is not essential for perceiving spoken language” (p. 301) while Lotto, Hickok, and Holt (2009) posit that “there is no need to think that this interaction [with the production system] would be required of normal speech perception” (p. 3). In other words, we are left with a model of speech perception in the brain in which the motor