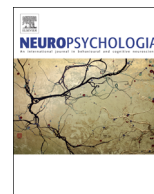




ELSEVIER

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Magnetic resonance imaging of the brain and vocal tract: Applications to the study of speech production and language learning



Daniel Carey, Carolyn McGettigan*

Department of Psychology, Royal Holloway, University of London, Egham, UK

ARTICLE INFO

Article history:

Received 27 January 2016

Received in revised form

2 June 2016

Accepted 5 June 2016

Available online 7 June 2016

Keywords:

Articulation

Functional MRI

Phonetic learning

Individual differences

Representational similarity analyses

ABSTRACT

The human vocal system is highly plastic, allowing for the flexible expression of language, mood and intentions. However, this plasticity is not stable throughout the life span, and it is well documented that adult learners encounter greater difficulty than children in acquiring the sounds of foreign languages. Researchers have used magnetic resonance imaging (MRI) to interrogate the neural substrates of vocal imitation and learning, and the correlates of individual differences in phonetic “talent”. In parallel, a growing body of work using MR technology to directly image the vocal tract in real time during speech has offered primarily descriptive accounts of phonetic variation within and across languages. In this paper, we review the contribution of neural MRI to our understanding of vocal learning, and give an overview of vocal tract imaging and its potential to inform the field. We propose methods by which our understanding of speech production and learning could be advanced through the combined measurement of articulation and brain activity using MRI – specifically, we describe a novel paradigm, developed in our laboratory, that uses both MRI techniques to for the first time map directly between neural, articulatory and acoustic data in the investigation of vocalisation. This non-invasive, multimodal imaging method could be used to track central and peripheral correlates of spoken language learning, and speech recovery in clinical settings, as well as provide insights into potential sites for targeted neural interventions.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

The sensorimotor plasticity of the human brain is essential to the acquisition of spoken language. When learning a second language, or L2, it has been shown that the age at which learning begins has a substantial impact on how native or natural spoken pronunciation of that language sounds (Flege et al., 1999, 1995). An increased knowledge of learning in the adult vocal system is, in general, critical to the wider understanding of human communicative behavior, from the flexibility of self-expression in conversation to the recovery of speech and functional reorganization after brain injury. However, speech is one of the most complex actions we perform, with equally complex acoustic consequences. In order to understand its mechanistic underpinnings, we must identify adequate methodology with which to measure and link central neural processes, the actions of peripheral effector systems (i.e. the larynx and articulators) and the acoustic correlates of speech. Magnetic resonance imaging (MRI) offers the opportunity to achieve such a comprehensive account of vocal behavior. In this review, we examine the neuroimaging literature on vocal learning

and sensorimotor adaptation, including individual differences in these processes, and point to some of the challenges of assessing speech performance. We provide an introduction to the MRI of vocal tract dynamics as a means of obtaining performance measures more proximal to the motor task of speech production. We outline the applications of vocal tract imaging techniques to date, and propose how these can be incorporated into cognitive neuroscience via a methodological approach in which acoustic, articulatory and neural data can be integrated via analyses of representational similarities.

2. Phonetic learning in the brain: Functional and structural underpinnings identified with MRI

Humans are vocal learners, with a sophisticated capacity to volitionally inflect speech and vocalizations dependent on acoustic, linguistic and social contexts (McGettigan, 2015; Pisanski et al., 2016). Imitation of heard speech is largely instinctive to infant language learners, and as a task, is readily achievable for adult spoken language users – so much so that convergence on spoken pronunciation can occur in the absence of awareness (Kappes et al., 2009; Pardo, 2006a; Pardo et al., 2012; Pardo and Jay, 2010,

* Corresponding author.

Pardo, 2006b). As part of the acquisition of a spoken language, the learner must engage in the rehearsal of spoken material, demanding the sensorimotor transformation of heard (or imagined) signals into motor plans for execution. Studies of this basic process, and variability within it, thus bear great relevance for our understanding of phonetic learning.

Although many functional Magnetic Resonance Imaging (fMRI) studies have explored the neural correlates of language production, including investigations of bilingualism and language switching, relatively few have specifically addressed components pertaining to phonetic learning and flexibility. In general, these have reported variation in the structure and function of parts of the speech perception and production networks, related to differences in the production of non-native speech sounds due to learning/training, or individual variability in imitative skill (in monolingual and multilingual talkers). Here, we make distinct those studies that have explored phonetic learning in terms of speech perception and perceptual category formation from those that have focused on the imitation of perceptually discriminable stimuli (and hence are weighted toward audio-motor transformation and execution of speech). In the following sections, we restrict our discussion to the latter category.

2.1. Functional neuroimaging studies of phonetic learning and imitation

The performance of overt speech imitation and phonological/phonetic manipulations typically involves the flexible manipulation of vocal output, including (but not limited to) the phonemes produced, the language engaged (i.e., L1 vs. L2), and the properties of the voice (e.g., pitch, accent, formant spacing). Neuroimaging studies consistently implicate regions of fronto-parietal cortex in these processes, typically including the opercular part of the left inferior frontal gyrus and the inferior parietal cortex (in particular, supramarginal gyrus; Golestani and Pallier, 2007; Moser et al., 2009; Peschke et al., 2012; Reiterer et al. 2013, 2011; Simmonds et al., 2011), as well as the anterior insula (McGettigan et al., 2013; Moser et al., 2009). These regions are major nodes in the dorsal pathway, described in several models of speech processing as having an important role in sensorimotor transformations for speech comprehension and production (Hickok and Poeppel, 2007; Rauschecker and Scott, 2009; Scott and Johnsrude, 2003). Models of speech production implicate inferior frontal regions and adjacent premotor cortex in the representation of motor representations for learned and familiar articulations (ranging from phonemes to syllables and wordforms, depending on the model; Guenther and Vladusich, 2012; Hickok, 2012; Hickok et al., 2011; Rauschecker and Scott, 2009; Tourville and Guenther, 2011). Most views posit that speech production is supported by a set of internal feedforward and feedback models, in which the system acts to minimize output errors by comparing the predicted and actual outcomes of speech in auditory and somatosensory cortices. The specific role of inferior parietal cortex varies slightly, with Guenther and colleagues proposing a direct somatosensory processing function, where Rauschecker and Scott (2009) suggest a more general “hub”-like role in matching feedforward signals from inferior frontal sites with sensory feedback signals from posterior temporal cortex. As inferior parietal cortex has been proposed as a putative substrate for the phonological store (Buchsbaum and D’Esposito, 2008; Jacquemot et al., 2003; Jacquemot and Scott, 2006), this is also considered as a possible role for areas such as the supramarginal gyrus in phonetic imitation (Reiterer et al., 2013, 2011). Several of the studies described in the coming sections report activation of the insula during imitation and learning. In the context of clinical speech pathology, a seminal lesion overlap study of apraxia of speech by Dronkers (1996) argued for a

crucial role of the tip of the left precentral gyrus of the insula in the motor control of speech – however, its implication in functional neuroimaging studies of speech perception and production, and of language input and output processes more generally, has led others to suggest that the insula may act as a hub for information from speech and language processing areas (Oh et al., 2014).

Voluntary, on-demand changes in the sound of articulated speech – for example sounding more masculine or feminine (Cartei et al., 2012), or conveying particular personality traits (Hughes et al., 2014) – are relatively intuitive for human talkers. These behaviours can provide insights into the plasticity of speech systems (Pisanski et al., 2016). McGettigan et al. (2013) asked participants to perform spoken impressions of known talkers (from visual prompts) and found, in a comparison of altered versus normal speech, that the left anterior insula and pars opercularis were commonly engaged to emulate general accents/styles of speaking and specific vocal identities. In a similar study, participants performing overt phonological manipulations (i.e. modifying the prosodic or segmental content by producing inflected versions of heard words) showed increased activation of the inferior frontal gyrus (pars triangularis, and extending onto pars opercularis for segmental manipulations only) and the intraparietal sulcus (Peschke et al., 2012).

Garnier and colleagues aimed to investigate the difference between conscious imitation and unconscious phonetic convergence effects in the overt repetition of heard vowels (Garnier et al., 2013). They reported that participants' speech showed significant correlations with target pitch for both overt imitation and phonetic convergence. In an ROI analysis of functional MRI responses, they further reported that the degree of acoustic imitation shown in an individual was positively related to activation of bilateral auditory cortex, left Wernicke's area, and bilateral inferior parietal regions including supramarginal gyrus. Another study found that the degree of participants' relatively unconscious matching to the duration of targets in a rapid speech shadowing task was correlated with activation in the right inferior parietal cortex (in the region of the supramarginal gyrus; Peschke et al., 2009).

Several studies have more directly modeled language learning and proficiency by testing imitation of non-native speech, and measuring time- and learning-dependent changes in neural activity (Moser et al., 2009; Segawa et al., 2013; Simmonds et al., 2014). Moser et al. (2009) found overall increases in the neural response for imitation of heard non-native > English nonwords, where ROI analyses of the left inferior frontal gyrus and anterior insula showed marginal and significant decreases in activation throughout the course of the experiment, respectively; notably, an individual differences analysis revealed a significant correlation between the response of the left anterior insula and the amount of behavioural improvement during the scan. A study investigating the learning-related changes in the overt pronunciation of phonotactically illegal pseudowords in English (e.g. GVAZF) compared reading aloud of trained and novel items in fMRI (Segawa et al., 2013). Here, novel items generated greater activation in both frontal opercula, supplementary motor area (SMA), left superior temporal cortex and bilateral superior parietal lobule, as well as the globus pallidus. Simmonds et al. (2015) scanned participants before and after an intensive 1-week self-administered programme of foreign language pronunciation training. They reported a significant increase in the BOLD response during imitation of non-native (vs native) speech, in bilateral auditory and motor cortices, the basal ganglia, thalamus and cerebellum. However, while these authors also report activations in bilateral frontal operculum and inferior parietal lobule for this contrast, they attribute this engagement not to sensorimotor processes but rather to cingulo-opercular salience and fronto-parietal central executive

Download English Version:

<https://daneshyari.com/en/article/5045282>

Download Persian Version:

<https://daneshyari.com/article/5045282>

[Daneshyari.com](https://daneshyari.com)