Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research report

Mental imagery of speech implicates two mechanisms of perceptual reactivation

Xing Tian [a,b,c,*], *Jean Mary Zarate* [c] and *David Poeppel* [c,d]

[a] *New York University Shanghai, China*
[b] *NYU-ECNU Institute of Brain and Cognitive Science at NYU Shanghai, China*
[c] *Department of Psychology, New York University, USA*
[d] *Department of Neuroscience, Max Planck Institute for Empirical Aesthetics (MPIEA), Germany*

ARTICLE INFO

Article history:
Received 3 February 2015
Reviewed 18 March 2015
Revised 26 June 2015
Accepted 5 January 2016
Action editor Alessandro Tavano
Published online 14 January 2016

Keywords:
Prediction
Mental simulation
Sensorimotor integration
Internal forward model/efference copy/corollary discharge
Memory retrieval

ABSTRACT

Sensory cortices can be activated without any external stimuli. Yet, it is still unclear how this perceptual reactivation occurs and which neural structures mediate this reconstruction process. In this study, we employed fMRI with mental imagery paradigms to investigate the neural networks involved in perceptual reactivation. Subjects performed two speech imagery tasks: articulation imagery (AI) and hearing imagery (HI). We found that AI induced greater activity in frontal-parietal sensorimotor systems, including sensorimotor cortex, subcentral (BA 43), middle frontal cortex (BA 46) and parietal operculum (PO), whereas HI showed stronger activation in regions that have been implicated in memory retrieval: middle frontal (BA 8), inferior parietal cortex and intraparietal sulcus. Moreover, posterior superior temporal sulcus (pSTS) and anterior superior temporal gyrus (aSTG) was activated more in AI compared with HI, suggesting that covert motor processes induced stronger perceptual reactivation in the auditory cortices. These results suggest that motor-to-perceptual transformation and memory retrieval act as two complementary mechanisms to internally reconstruct corresponding perceptual outcomes. These two mechanisms can serve as a neurocomputational foundation for predicting perceptual changes, either via a previously learned relationship between actions and their perceptual consequences or via stored perceptual experiences of stimulus and episodic or contextual regularity.

1. Introduction

Sensory cortices can be activated without any external stimulation (e.g., Ji & Wilson, 2006; Wheeler, Petersen, & Buckner, 2000). That is, perceptual neural representations can be reconstructed without perceptual processing (referred to as perceptual reactivation). Mental imagery, defined as an internally generated quasi-perceptual experience, is one such example (e.g., Kosslyn et al., 1999; Kraemer, Macrae, Green, & Kelley, 2005). The ability to form mental images has been

* *Corresponding author*. New York University Shanghai 1555 Century Avenue, Room 1138 Pudong, Shanghai 200122, China.
E-mail address: xing.tian@nyu.edu (X. Tian).
http://dx.doi.org/10.1016/j.cortex.2016.01.002

hypothesized as a vehicle for generating and representing thoughts. This argument can be found as early as Plato's Theaetetus [427–347 BC] (1987) and Aristotle's De Anima [384–322 BC] (1986). In the age of enlightenment, mental imagery was considered analogous to perception by philosophers such as Descartes (1642/1984), Hobbes (1651/1968), Berkeley (1734/1965a, 1734/1965b) and Hume (1969). Early experimental psychologists such as Wundt (1913) and James (1890) proposed that ideas were represented as mental images in both visual and auditory domains. Modern research in mental imagery has yielded insight on how thought is represented in cognitive systems (Kosslyn, 1994; Kosslyn, Ganis, & Thompson, 2001; Paivio, 1971, 1986; Pylyshyn, 1981, 2003).

Recently, an additional computational role of mental imagery has been proposed: a mechanism to plan possible future contingencies. That is, mental imagery has been modeled as a process in which perceptual consequences can be predicted to gain advantages in various aspects of perception, memory, decision making and motor control (Albright, 2012; Moulton & Kosslyn, 2009; Schacter et al., 2012; Tian & Poeppel, 2012). The reactivation of perceptual neural representations without any external stimulation is the key mechanism mediating this predictive ability (Moulton & Kosslyn, 2009). Internally induced neural representations, which are highly similar to the ones established in corresponding perceptual processing, have been observed in modality-specific areas, such as in visual (e.g., Kosslyn et al., 1999; O'Craven & Kanwisher, 2000), auditory (e.g., Kraemer et al., 2005; Shergill et al., 2001; Zatorre, Halpern, Perry, Meyer, & Evans, 1996), somatosensory (e.g., Yoo, Freeman, McCarthy III, & Jolesz, 2003; Zhang, Weisser, Stilla, Prather, & Sathian, 2004) and olfactory (e.g., Bensafi et al., 2003; Djordjevic, Zatorre, Petrides, Boyle, & Jones-Gotman, 2005) domains.

It is not clear how these neural representations are reconstructed. Preliminary evidence from an MEG study (Tian & Poeppel, 2013) suggests that imagining speaking (articulation imagery, AI) and imagining hearing (hearing imagery, HI) differentially modulated neural responses to subsequent auditory stimuli. These distinct modulation effects by different types of imagery suggest that similar auditory neural representations may be internally formed via different neural pathways. A dual stream prediction model (DSPM, Fig. 1) was proposed in which two distinct processes in parallel neural pathways can internally induce the corresponding perceptual neural representation (Tian & Poeppel, 2012, 2013).

In the *simulation-estimation prediction stream* (Fig. 1), the perceptual consequences of actions are predicted by simulating the movement trajectory, followed by estimating the perceptual changes that would be associated with this movement. AI has been hypothesized to implement the motor-to-sensory transformation for simulation-estimation mechanism (Tian & Poeppel, 2013). Specifically, during AI, a motor simulation process similar to speech motor preparation is carried out, but without execution and output (Palmer et al., 2001; Tian & Poeppel, 2010, 2012). Therefore, neural networks that mediate motor simulation should be similar to the ones implicated in motor preparation, including supplementary motor area (SMA), inferior frontal gyrus (IFG), premotor and insula (Bohland & Guenther, 2006; Palmer et al., 2001; Shuster & Lemieux, 2005). After motor simulation, a copy of the planned motor commands – known as the efference copy (Von Holst & Mittelstaedt, 1950/1973; for a review see Wolpert & Ghahramani, 2000) – is sent to the somatosensory areas and is used in a forward model to estimate the somatosensory consequences (Blakemore & Decety, 2001). This somatosensory estimation is hypothesized to be governed by the networks underlying somatosensory perception (Blakemore, Wolpert, & Frith, 1998; Tian & Poeppel, 2010, 2012), including primary and secondary somatosensory regions, parietal operculum (PO) and the supramarginal gyrus (SMG). Moreover in the context of speech, we hypothesize that auditory consequences are predicted on the basis of somatosensory estimation, and this auditory estimation will recruit neural structures in temporal auditory cortices (Tian & Poeppel, 2010, 2012, 2013, 2015).

In the *memory-retrieval prediction stream* (Fig. 1), the internally induced neural representations are the result of memory retrieval processes – reconstructing stored perceptual information in modality-specific cortices (Kosslyn, 1994, 2005; Wheeler et al., 2000). In particular, the retrieved object properties from long-term memory reactivate the sensory cortices that originally processed the object features (Kosslyn, 1994). In this experiment, we employed HI to probe this memory-retrieval stream. Auditory representations can be retrieved from various memory sources such as episodic memory, which presumably relies on hippocampal structures (Carr, Jadhav, & Frank, 2011; Eichenbaum, Sauvage, Fortin, Komorowski, & Lipton, 2012) with a possible buffer site in parietal cortex (Vilberg & Rugg, 2008; Wagner, Shannon, Kahn, & Buckner, 2005). Auditory representations can also be transformed from lexical and semantic information stored in semantic networks, including frontal (e.g., dorsomedial prefrontal cortex, IFG, ventromedial prefrontal cortex), parietal (e.g., posterior inferior parietal lobe) and temporal (e.g., middle temporal gyrus) regions (Binder, Desai, Graves, & Conant, 2009; Lau, Phillips, & Poeppel, 2008; Price, 2012). Regardless of the divergent functional roles (episodic or semantic networks), frontal and parietal regions are reliably activated during memory retrieval processes. Therefore, neural activation in a frontal-parietal distributed network – the proposed memory-retrieval prediction stream – should be observed during HI.

This study uses fMRI to investigate three neuroanatomical/functional hypotheses that are generated from the DSPM. First, if the perfect simulation-estimation and memory-retrieval tasks were carried out, two distinct processing streams would be revealed separately. However, because speech imagery could involve both production and perception, we predict that both types of imagery will activate the simulation-estimation stream for simulating speech motor action (Tian & Poeppel, 2013). More importantly, we hypothesize that each type of imagery will recruit each prediction stream to a different extent. Specifically, we predict that AI will induce stronger activation in the simulation-estimation prediction stream, including SMA, IFG, premotor, insula for motor simulation, as well as primary and/or secondary somatosensory areas PO and SMG for subsequent estimation of somatosensory consequences. On the other hand, we predict that HI will have more activation in the memory-retrieval prediction stream, which is comprised of frontal, superior