NeuroImage 142 (2016) 565–575

Contents lists available at ScienceDirect

NeuroImage

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

Regional homogeneity of intrinsic brain activity correlates with auditory-motor processing of vocal pitch errors

Zhiqiang Guo [a,1], Xiyan Huang [b,1], Meng Wang [c,1], Jeffery A. Jones [d], Zhengjia Dai [e], Weifeng Li [b], Peng Liu [b], Hanjun Liu [b,f,*]

[a] Department of Biomedical Engineering, School of Engineering, Sun Yat-sen University, Guangzhou 510006, China
[b] Department of Rehabilitation Medicine, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou 510080, China
[c] Department of Radiology, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou 510080, China
[d] Psychology Department and Laurier Centre for Cognitive Neuroscience, Wilfrid Laurier University, Waterloo, Ontario, N2L 3C5, Canada
[e] Department of Psychology, Sun Yat-sen University, Guangzhou 510006, China
[f] Guangdong Provincial Key Laboratory of Brain Function and Disease, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou 510080, China

ARTICLE INFO

Article history:
Received 11 May 2016
Accepted 3 August 2016
Available online 5 August 2016

Keywords:
Auditory feedback
Intrinsic brain activity
Resting-state fMRI
Auditory-vocal integration
Regional homogeneity

ABSTRACT

It has been well documented that speakers produce rapid compensatory vocal adjustments for errors they perceive in their auditory feedback. The fact that they differ greatly in the degree to which they compensate for perceived errors, however, has received much less attention. The present study investigated whether intrinsic brain activity during resting can predict an individual's behavioral and cortical responses in compensating for pitch-shifted auditory feedback during vocalization. This relationship was investigated by correlating the regional homogeneity (ReHo) of resting-state fMRI signals with the vocal compensation and event-related potentials (N1 and P2) in response to pitch shifts of −200 and −500 cents. Behaviorally, the magnitudes of vocal compensation were significantly correlated with the ReHo values in the right supplementary motor area (SMA) for both −200 and −500 cents, the right primary motor cortex (M1) for −200 cents, and the left premotor cortex (PMC) for −500 cents. For both pitch shift sizes, there were significant correlations between ReHo and N1 amplitude in the left inferior frontal gyrus (IFG), right superior temporal gyrus (STG), bilateral M1, and left SMA. Significant correlations between ReHo and P2 amplitude were observed in the bilateral IFG, right STG, left SMA and M1 for −200 and −500 cents, the left PMC for −200 cents, and the right SMA for −500 cents. These findings provide the first evidence that regional homogeneity of intrinsic brain activity can predict behavioral and cortical responses in compensating for pitch errors in voice auditory feedback.

1. Introduction

Speech is one of the most complex motor behaviors produced by humans. The control of speech involves the integration of sensory information into the vocal motor systems (Smotherman 2007). Although somatosensory feedback is important (Lametti et al. 2012), auditory feedback has been demonstrated to be particularly important for monitoring and correcting errors in speech production (Houde and Jordan 1998; Jones and Munhall 2005; Perkell 2012). For example, speakers produce rapid compensatory vocal adjustments when they hear mismatches between their auditory feedback and their intended vocal fundamental frequency (F_0), (Burnett et al. 1998; Jones and Munhall 2002), intensity (Bauer et al. 2006; Liu et al. 2007), or formant frequency (Houde and Jordan 1998; Purcell and Munhall 2006). This compensatory process is thought to stabilize the production of speech around the desired speech target and is guided by task demands (Chen et al. 2007).

To understand the sensorimotor integration that underlies voice production, researchers have investigated the neural correlates of vocal compensation for pitch feedback errors by exposing speakers to frequency-altered auditory feedback (FAF). For example, event-related potential (ERP) studies have shown that the N1-P2 complex elicited by FAF is modified by stimulus features (Liu et al., 2011a; Scheerer et al., 2013a), physiological development (Liu et al. 2013; Scheerer et al., 2013b), attentional demands (Hu et al. 2015; Liu et al. 2015), cognitive training (Chen et al. 2015; Li et al. 2015), and the nature of the vocalization task (Behroozmand et al. 2009; Behroozmand et al. 2011). Studies using functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), and electrocorticography (ECoG) have shown that the brain regions involved in auditory-motor integration for voice control include the superior temporal gyrus (STG), primary motor cortex (M1), superior temporal sulcus (STS), anterior cingulate cortex

* Corresponding author at: Department of Rehabilitation Medicine, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou 510080, China.
E-mail address: lhanjun@mail.sysu.edu.cn (H. Liu).
[1] Authors contributed equally.

http://dx.doi.org/10.1016/j.neuroimage.2016.08.005

(ACC), premotor cortex (PMC), supplementary motor area (SMA), inferior frontal gyrus (IFG), and the anterior insula (Behroozmand et al. 2015; Chang et al. 2013; Greenlee et al. 2013; Kort et al. 2014; Parkinson et al. 2012; Toyomura et al. 2007; Zarate et al. 2010; Zarate and Zatorre 2008). Activity in the auditory (e.g. STG) and motor (e.g. PMC) cortices has been found to significantly correlate with the trial-to-trial magnitude of the vocal compensations for pitch-shifted auditory feedback within subjects (Chang et al. 2013) and the mean compensation magnitudes across subjects (Behroozmand et al. 2015).

Despite the progress made in understanding the neural mechanisms involved in auditory-motor integration for voice control, one important issue that has not received sufficient attention is the large variability of the magnitude of vocal compensation, or the pitch-shift reflex (PSR), for vocal errors heard in auditory feedback observed across individuals (Burnett et al. 1998; Chen et al. 2013; Liu et al., 2011b; Natke and Kalveram 2001; Scheerer and Jones 2012). For example, when exposed to auditory feedback that deviated from their true production by 25–300 cents, vocal compensations have been reported to range between 3 and 100 cents for healthy young adults (Burnett et al. 1998) and between 10 and 107 cents for individuals with Parkinson's disease (PD) (Chen et al. 2013). Similarly, large inter-individual variability in the N1 and P2 responses to FAF has also been observed in several ERP studies (Behroozmand et al. 2014; Chen et al. 2012; Li et al. 2013; Scheerer et al., 2013b). Despite the significant inter-individual differences in the production of vocal compensation for FAF, most previous studies have focused on the mean differences observable between tasks and groups. This exclusive emphasis on measures of central tendency may obscure important individual differences that could shed light on the mechanisms that support the integration of sensory and motor information that is required for control of such a complex motor behavior (MacDonald et al. 2006).

The few studies that have explored the individual characteristics that might explain the difference in the production of vocal compensation for FAF suggest that important differences lie on both sides of the arbitrary line between the motor and sensory systems. For example, the variability of the speaker's normal (i.e. unaltered) voice F_0 is predictive of the magnitude of vocal compensation for FAF (Chen et al. 2013; Scheerer and Jones 2012). As well, a significant negative correlation was found between the latency of the N1 response to FAF and the Montreal Battery of Evaluation of Amusia (MBEA) scores (Korzyukov et al. 2012), which suggests that people who have better musical perception process pitch feedback errors during vocalization more rapidly in the cortex. Although these studies suggest certain individual differences can explain part of the variability in cortical and behavioral processing of feedback errors, very little is known about the neural underpinnings that determine an individual's predisposition for auditory-motor integration for voice.

A number of studies have shown that intrinsic brain activity evaluated during resting-state can predict inter-individual variability in a wide range of behavioral tasks and cognitive functions. For example, resting-state fMRI has revealed intrinsic neural networks that support and overlap with functional networks recruited during visual, auditory, language, and salience detection tasks (Biswal et al. 1995; Fox et al. 2006; Fox et al. 2005; Hampson et al. 2002; Smith et al. 2009). Moreover, resting-state measures have been shown to correlate with inter-individual performance variability across several cognitive domains (Hampson et al. 2006; Koyama et al. 2011; Seeley et al. 2007; Wang et al. 2014; Zou et al. 2013). For example, working memory performance (n-back) can be predicted by the functional connectivity between the medial frontal gryus (MFG) and ventral anterior cingulate cortex (vACC) to posterior cingulate cortex (PCC) in the resting condition (Hampson et al. 2006) and the regional amplitude of low-frequency fluctuation (ALFF) of resting-state activity in the superior parietal lobule/precuneus (Zou et al. 2013). In addition, regional homogeneity (ReHo) in the DLPFC can predict the individual variability of performance on a task of the behavioral conflict adaptation (Wang et al. 2014).

In the context of speech processing, however, the relationship between intrinsic brain activity and behavioral performance is rarely reported. In a recent resting-state functional connectivity study investigating the neural networks associated with the control of vocalization in PD, individuals with PD who experienced more difficulty with speech communication (item 5 of UPDRS-II) exhibited decreased connectivity between the right ventral PMC and right STG, the left ventral PMC and right putamen, and between the right putamen and left thalamus (New et al. 2015). Whether resting-state measures are predictive of individual differences in the feedback-based online monitoring of vocal production, however, remains unknown.

The present study was thus designed to investigate whether intrinsic brain activity assessed by resting-state fMRI can predict an individual's performance in auditory-motor integration for voice control behaviorally and neurally. Participants were instructed to sustain a vowel phonation while hearing their voice auditory feedback unexpectedly pitch-shifted in real time. The magnitudes and latencies of vocal and ERP responses (N1 and P2) in the production of compensation for FAF were measured and correlated with intrinsic brain activity indexed by ReHo (Zang et al. 2004). ReHo describes the summarized local functional synchronization of a given voxel to its nearest neighboring nodes, and it has been used to investigate the neural basis of individual differences in behavioral performance (Tian et al. 2012; Wang et al. 2011; Zou et al. 2013). We measured the ReHo values within a network of brain regions that was previously identified and showed robust activation in auditory-motor integration for voice control, including the STG, M1, PMC, SMA and IFG (Behroozmand et al. 2015; Parkinson et al. 2012). We expected that inter-individual variability in the production of vocal compensation for FAF at the levels of behavior and cortex would be associated with the ReHo in the network of brain regions involved in the control of vocalization.

2. Materials and methods

2.1. Subjects

Eighteen right-handed native Mandarin Chinese speakers (17 female, age: 21.1 ± 1.7 years) from Sun Yat-sen University of China participated in this experiment. They reported no history of language, hearing, or neurological disorders. All participants passed a hearing screen with hearing thresholds of ≤25 dB HL at 250–4000 Hz, and gave informed consent and received monetary compensation for their participation. The experiment was carried out following the research protocol approved by the Institutional Review Board of The First Affiliated Hospital at Sun Yat-sen University of China in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

2.2. MRI scanning

MRI data were obtained during a resting state using a Siemens Magnentom 3T Trio Tim MRI scanner (Erlangen, Germany) located in the Brain Imaging Center at South China Normal University. Each participant was instructed to lie still with their eyes open and think of nothing in particular. Functional images were obtained using an echo-planar imaging (EPI) pulse sequence: echo time (TE) = 30 ms; repetition time (TR) = 2000 ms; slice thickness = 3.5 mm; voxel size = 1 × 1 × 1 mm; flip angle = 90°; field of view (FOV) = 224 × 224 mm^2. For registration purposes, a set of anatomical images was obtained through a T1-weighted sequence: TR = 2300 ms; TE = 3.24 ms; slices thickness = 1 mm; voxel size = 1 × 1 × 1 mm; flip angle = 9°; FOV = 256 × 256 mm^2. The scanning lasted for 8 min and produced 240 brain volumes.