NeuroImage xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

NeuroImage

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

Sensory–motor networks involved in speech production and motor control: An fMRI study

Roozbeh Behroozmand [a,b,*], Rachel Shebek [a], Daniel R. Hansen [a], Hiroyuki Oya [a], Donald A. Robin [c], Matthew A. Howard III [a], Jeremy D.W. Greenlee [a]

[a] Human Brain Research Lab, Department of Neurosurgery, University of Iowa, Iowa City, IA 52242, United States
[b] Speech Neuroscience Lab, Department of Communication Sciences and Disorders, University of South Carolina, Columbia, SC 29208, United States
[c] Research Imaging Institute, Departments of Neurology, Radiology and Biomedical Engineering, University of Texas Health Science Center San Antonio, San Antonio, TX 78229, United States

ARTICLE INFO

Article history:
Accepted 17 January 2015
Available online xxxx

Keywords:
Speech motor control
Auditory feedback
Sensory–motor integration
Pitch perturbation
fMRI

ABSTRACT

Speaking is one of the most complex motor behaviors developed to facilitate human communication. The underlying neural mechanisms of speech involve sensory–motor interactions that incorporate feedback information for online monitoring and control of produced speech sounds. In the present study, we adopted an auditory feedback pitch perturbation paradigm and combined it with functional magnetic resonance imaging (fMRI) recordings in order to identify brain areas involved in speech production and motor control. Subjects underwent fMRI scanning while they produced a steady vowel sound /a/ (speaking) or listened to the playback of their own vowel production (playback). During each condition, the auditory feedback from vowel production was either normal (no perturbation) or perturbed by an upward (+600 cents) pitch-shift stimulus randomly. Analysis of BOLD responses during speaking (with and without shift) vs. rest revealed activation of a complex network including bilateral superior temporal gyrus (STG), Heschl's gyrus, precentral gyrus, supplementary motor area (SMA), Rolandic operculum, postcentral gyrus and right inferior frontal gyrus (IFG). Performance correlation analysis showed that the subjects produced compensatory vocal responses that significantly correlated with BOLD response increases in bilateral STG and left precentral gyrus. However, during playback, the activation network was limited to cortical auditory areas including bilateral STG and Heschl's gyrus. Moreover, the contrast between speaking vs. playback highlighted a distinct functional network that included bilateral precentral gyrus, SMA, IFG, postcentral gyrus and insula. These findings suggest that speech motor control involves feedback error detection in sensory (e.g. auditory) cortices that subsequently activate motor-related areas for the adjustment of speech parameters during speaking.

Introduction

During speaking, the brain coordinates the movement of respiratory, laryngeal, articulatory and facial muscles in order to produce speech sounds. This task requires the involvement of feedforward mechanisms that mediate speech production and motor control for effective communication (Guenther et al., 2006). However, questions remain as to how the brain monitors speech production to ensure performance accuracy. Evidence provided by several studies shows that sensory feedback information (e.g. auditory and somatosensory) plays a critical role during speech production (Houde, 1998; Lametti et al., 2012; Larson, 1998). The brain continuously monitors feedback information in order to correct for unwanted production errors and update the state of the sensory–motor networks to accomplish current and future speech production goals. These critically important networks and the underlying neural mechanisms that incorporate sensory feedback to optimize human speech motor behavior are poorly understood.

A well-accepted theory has proposed that the brain manages to produce and monitor speech by comparing the incoming sensory feedback information with an internal representation of the predicted feedback (Hickok et al., 2011; Rauschecker and Scott, 2009). These internal predictions are hypothesized to be generated by an internal forward model (Wolpert et al., 2011) that transmits efference copies of the speech motor commands to sensory modalities in order to characterize and detect disparities (errors) between intended and actual speech feedback. In case of a mismatch between the predicted and actual sensory feedback information, the output of this comparative process will result in generation of an error signal that is projected back from the sensory to motor systems such that speech motor parameters are adjusted to improve production accuracy.

A widely-used experimental strategy to examine the interactions between sensory–motor mechanisms of speech is to apply a perturbation to the auditory feedback while human subjects speak. This

* Corresponding author at: University of South Carolina, Department of Communication Sciences and Disorders, Keenan Building, Room 356, 1224 Sumter St., Columbia, SC 29208, United States. Fax: +1 803 777 3081.
E-mail address: r-behroozmand@sc.edu (R. Behroozmand).

http://dx.doi.org/10.1016/j.neuroimage.2015.01.040
1053-8119/

technique allows experimenters to externally induce a mismatch between internally-predicted and actual sensory feedback information to understand how the brain detects feedback errors and uses them for speech production and motor control. From a behavioral standpoint, studies have shown that auditory feedback perturbation elicits compensatory vocal reactions that changes speech parameters (e.g. pitch, formant or loudness) in the opposite direction of the applied perturbation (Cai et al., 2011; Chen et al., 2007; Larson, 1998; Villacorta et al., 2007). This evidence demonstrates that the brain continuously monitors speech and operates like a feedback-based controller that uses auditory information for speech motor control.

The neural bases of such sensory–motor mechanisms and behavioral compensation have been studied by quantifying the neurophysiological correlates (e.g. EEG, MEG, ECoG) of speech production and motor control in an auditory feedback perturbation paradigm (Behroozmand and Larson, 2011; Behroozmand et al., 2009, 2011; Chang et al., 2013; Flinker et al., 2010; Greenlee et al., 2011, 2013; Heinks-Maldonado et al., 2005, 2006; Houde et al., 2002; Sitek et al., 2013). Results of these studies showed that the motor act of speaking modulates speech sound processing in auditory cortical areas. This modulatory effect was examined by comparing neural responses to perturbed speech feedback during both speaking and passive listening to the playback of self-produced speech. The major findings indicated that neural activity in auditory cortex was largely suppressed during speaking compared with playback conditions (Behroozmand and Larson, 2011; Chang et al., 2013; Flinker et al., 2010; Greenlee et al., 2011; Heinks-Maldonado et al., 2005, 2006; Houde et al., 2002; Sitek et al., 2013) and suppression was maximum when predicted and actual speech feedback were closely matched (i.e. no or small feedback error) (Behroozmand and Larson, 2011). However, when feedback was briefly perturbed in the middle of speech, neural responses within auditory cortex were increased during speaking (Behroozmand et al., 2009; Chang et al., 2013; Greenlee et al., 2013; Liu et al., 2010). It is hypothesized that such response modulation underlies detection of feedback changes and correction of unintended speech errors during speaking. Studies in non-human primates using a similar experimental task during vocal production yielded similar results that were consistent with the findings in the human brain (Eliades and Wang, 2008).

A recent functional magnetic resonance imaging (fMRI) study adopted the feedback perturbation paradigm to identify brain areas involved in speech motor control using perturbed auditory feedback (Parkinson et al., 2012). In that study, Parkinson et al. (2012) examined the fMRI correlates of speech feedback processing during an active speaking task with and without pitch perturbation in the auditory feedback. Results revealed a complex sensory–motor network involved in speech feedback processing including superior temporal gyrus (STG), precentral gyrus, postcentral gyrus, supplementary motor area (SMA), inferior frontal gyrus (IFG), inferior parietal lobule (IPL) and insula. In the present study, we used a similar approach but with the addition of the playback condition under which subjects listened to the playback of their own perturbed and unperturbed speech feedback. This combined approach created new experimental contrasts for comparing speaking vs. playback conditions in order to isolate sensory–motor networks of speech and study them independently. Furthermore, since the contrast between perturbed and unperturbed auditory feedback for small perturbations (±100 cents) in Parkinson et al.'s study (Parkinson et al., 2012) did not yield significant differences, we increased the magnitude of the pitch perturbations to +600 cents in order to identify brain areas involved in feedback error processing during speech. It has previously been shown that increasing the pitch-shift stimulus magnitude elicits larger ERP responses, suggesting that the brain generates a larger error signal in response to increased degree of mismatch between one's own voice pitch and its auditory feedback (Behroozmand et al., 2009). Based on this evidence, we used a +600 cents pitch-shift stimulus to elicit greater activation of BOLD responses in order to improve signal to noise ratio compared to Parkinson et al.'s study (Parkinson et al., 2012). We hypothesized that larger pitch-shift stimulus magnitudes will enable us to highlight functional neural mechanisms of feedback error processing during vocal pitch monitoring and motor control.

Materials and methods

Subjects

We have previously used the feedback pitch perturbation paradigm in both speaking and playback conditions in human subjects undergoing the neurosurgical treatment of epilepsy. By recording electrocorticograms (ECoG) in those subjects, we have described auditory cortical responses after perturbed and unperturbed feedback states and reported neural response changes and the correlation of those changes with vocal behavior (i.e. compensation) in response to pitch shifts (Greenlee et al., 2013). Because ECoG provides limited anatomic sampling of only portions of an individual brain, we have begun using fMRI to provide greater anatomic sampling to supplement the insights gained from ECoG recordings.

Eight right-handed subjects (7 male and 1 female, mean age: 38 years) participated in this fMRI study days prior to surgery for subdural electrode implantation as part of a standard clinical treatment protocol for medically-intractable epilepsy. Formal neuropsychological testing was performed in all subjects before surgery and all had normal speech and language function. Audiometry was normal in all subjects. All subjects required pre-operative sodium amobarbital (i.e. Wada) testing of language dominance, and all were found to have left hemisphere dominance. The results of inpatient video-EEG monitoring demonstrated epileptic foci in the left hemisphere in five subjects, and in the right in three subjects. Patients were observed during fMRI scanning to ensure there were was no seizure events while they performed the experimental tasks in the scanner. All experimental procedures were approved by the University of Iowa Institutional Review Board.

Speech stimuli and experimental design

An event-related design was used to measure blood-oxygen-level dependent (BOLD) activation during speech with and without auditory feedback alteration. The experiment was carried out in one block during which subjects were instructed to either produce and maintain the steady vowel sound /a/ following the onset of a visual cue (speaking condition) or passively listen to the playback of their own self-production (playback condition). Subjects were instructed to hold their head still and minimize their movement during both speaking and playback tasks. Subjects' performance was monitored by the experimenters to ensure the minimal movement and consistency of the vocal production task during the whole recording session and proper feedback was given whenever it was necessary.

A visual cue projected onto a screen behind the scanner and visualized by the subject through a mirror was used to control the tasks and their timing. During speaking trials, the visual cue to speak consisted of a transition from a circle to a square icon. Subjects were instructed to continue their vowel production as long as the square was present on the screen (5 s). While the circle was presented, subjects were instructed to remain silent and listen to the playback of their own speech. An intermittent "Rest" condition was randomly included during which the circle was present on the screen and no speech was played back to the subjects. We chose to use simple circle and square cues to reduce the amount of BOLD signal activation resulting from the subjects reading written visual cues.

The speaking and playback conditions were interleaved in such a way that each speaking trial was recorded and immediately played back in the next trial (Fig. 1). Rest trials were randomly placed in between consecutive speaking-playback trials. During speaking trials, there was either no mismatch between speech and its auditory

Please cite this article as: Behroozmand, R., et al., Sensory–motor networks involved in speech production and motor control: An fMRI study, NeuroImage (2015), http://dx.doi.org/10.1016/j.neuroimage.2015.01.040