NeuroImage 55 (2011) 1324–1338

Contents lists available at ScienceDirect

NeuroImage

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

fMRI investigation of unexpected somatosensory feedback perturbation during speech

Elisa Golfinopoulos [a,*], Jason A. Tourville [a], Jason W. Bohland [a,b], Satrajit S. Ghosh [a,c], Alfonso Nieto-Castanon [d], Frank H. Guenther [a,c,e,f,g]

[a] Department of Cognitive and Neural Systems, Boston University, Boston, MA, USA
[b] Department of Health Sciences, Boston University, Boston, MA, USA
[c] Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA, USA
[d] StatsANC LLC, Buenos Aires, Argentina
[e] Department of Speech, Language, and Hearing Sciences, Boston University, Boston, MA, USA
[f] Division of Health Sciences and Technology, Harvard University–Massachusetts Institute of Technology, Cambridge, MA, USA
[g] Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, USA

ARTICLE INFO

Article history:
Received 3 June 2010
Revised 20 November 2010
Accepted 23 December 2010
Available online 30 December 2010

Keywords:
Functional magnetic resonance imaging
Structural equation modeling
Speech motor control
Right hemisphere
Somatosensory feedback

ABSTRACT

Somatosensory feedback plays a critical role in the coordination of articulator movements for speech production. In response to unexpected resistance to lip or jaw movements during speech, fluent speakers can use the difference between the somatosensory expectations of a speech sound and the actual somatosensory feedback to adjust the trajectories of functionally relevant but unimpeded articulators. In an effort to investigate the neural substrates underlying the somatosensory feedback control of speech, we used an event-related sparse sampling functional magnetic resonance imaging paradigm and a novel pneumatic device that unpredictably blocked subjects' jaw movements. In comparison to speech, *perturbed speech*, in which jaw perturbation prompted the generation of compensatory speech motor commands, demonstrated increased effects in bilateral ventral motor cortex, right-lateralized anterior supramarginal gyrus, inferior frontal gyrus *pars triangularis* and ventral premotor cortex, and bilateral inferior posterior cerebellum (lobule VIII). Structural equation modeling revealed a significant increased influence from left anterior supramarginal gyrus to right anterior supramarginal gyrus and from left anterior supramarginal gyrus to right ventral premotor cortex as well as a significant increased reciprocal influence between right ventral premotor cortex and right ventral motor cortex and right anterior supramarginal gyrus and right inferior frontal gyrus *pars triangularis* for *perturbed speech* relative to *speech*. These results suggest that bilateral anterior supramarginal gyrus, right inferior frontal gyrus *pars triangularis*, right ventral premotor and motor cortices are functionally coupled and influence speech motor output when somatosensory feedback is unexpectedly perturbed during speech production.

© 2011 Elsevier Inc. All rights reserved.

Introduction

Afferent feedback can be used during speech production to provide information about the position of an articulator relative to its target within a particular reference frame. This information can be used by the speaker to generate compensatory responses when the trajectory of a moving articulator is unexpectedly perturbed. Unexpected perturbation of lip or jaw movements during speech has been used extensively to investigate aspects of subjects' compensatory responses and probe the role of afferent feedback in speech motor control (e.g., Abbs and Gracco, 1984; Folkins and Abbs, 1975; Gomi et al., 2002; Gracco and Abbs, 1985; Ito et al., 2005; Kelso et al., 1984). These studies indicate that fluent adult speakers compensate for unexpected resistance to individual articulator movements by adjusting the kinematics of functionally relevant articulators outside the influence of the perturbation (e.g., Abbs and Gracco, 1984; Gomi et al., 2002; Gracco and Abbs, 1985; Ito et al., 2005; Kelso et al., 1984; Shaiman, 1989) and that speakers are capable of maintaining acoustic intelligibility in spite of dynamic perturbations to jaw movements (Nasir and Ostry, 2006, 2008, 2009; Tremblay et al., 2003). Several studies have also implicated a role for somatosensory feedback in the generation of compensatory responses to dynamic jaw perturbations (e.g., Nasir and Ostry, 2006, 2008, 2009; Tremblay et al., 2003). For example, Nasir and Ostry (2008) recently demonstrated that with their implants turned off cochlear implant subjects could produce compensatory responses to jaw perturbations that were comparable to normal-hearing control subjects, indicating that the generation of compensatory responses is possible in the complete absence of auditory feedback. Collectively these studies suggest that during habitual speech production when articulator movement is unexpectedly perturbed, the speech motor system can utilize somatosensory feedback to generate functionally equivalent articulatory gestures and compensate for the perturbation.

* Corresponding author. Department of Cognitive and Neural Systems, Boston University, 677 Beacon Street, Boston, MA 02215, USA. Fax: +1 617 353 7755.
E-mail address: egolfino@cns.bu.edu (E. Golfinopoulos).

1053-8119/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2010.12.065

Despite the large number of behavioral studies that have utilized static and dynamic perturbations to the articulators (e.g., Abbs and Gracco, 1984; Baum et al., 1996, 1997; Baum, 1999; Folkins and Abbs, 1975; Folkins and Zimmerman, 1982; Gay et al., 1981; Gomi et al., 2002; Gracco and Abbs, 1985; Jacks, 2008; Kelso et al., 1984; Lane et al., 2005; Lindblom et al., 1979; McFarland and Baum, 1995; Nasir and Ostry, 2006, 2008, 2009; Tremblay et al., 2003), few researchers have attempted to use non-invasive imaging techniques to investigate the neural substrates underlying the somatosensory feedback control of fluent speech production. Part of the challenge is that a device must be created that not only perturbs subjects' articulator movements, but is also MR compatible. In addition, the design must avoid employing a perturbation that would cause movement-related artifacts in the functional images. To perturb somatosensory feedback during functional magnetic resonance imaging (fMRI), Baciu et al. (2000) employed a tube fixed between subjects' lips during the articulation of a vowel that requires lip protrusion. Peak responses associated with the compensation were reported in bilateral cerebellum and dorsolateral prefrontal cortex, right supramarginal gyrus, right inferior frontal gyrus *pars triangularis*, right posterior superior temporal gyrus, and right middle temporal gyrus. These findings suggest that compensatory responses to static somatosensory perturbations during speech movements are mediated by a distributed network of brain regions that include cerebellar, temporal, and fronto-parietal areas and may generally involve greater contributions from right hemisphere cortical regions. However, due to the static nature of the perturbation in that study, speakers could have adapted their feedforward commands (or motor programs) to compensate. An unpredictable perturbation prompts subjects to rely more directly on the somatosensory feedback control subsystem (i.e., a subsystem that detects that somatosensory feedback is not within the expected range for the current speech sound and contributes to the adjustment of speech motor commands). Using transcranial magnetic stimulation, Ito et al. (2005) demonstrated that compensatory responses of the upper lip to unexpected perturbations of the jaw involved the left primary motor cortex. Since the study only targeted the primary motor cortex, the neural substrates underlying compensatory responses to unanticipated somatosensory perturbations during speech has yet to be fully characterized.

Insight into the possible neural and computational bases underlying somatosensory feedback control in speech comes from the DIVA (Directions into Velocities of Articulators) model of speech production and acquisition in which model cells are associated with neuroanatomical substrates (Golfinopoulos et al., 2010; Guenther, 1994, 1995, 2006; Guenther et al., 2006; Guenther and Vladusich, in press; Tourville and Guenther, 2010). According to the model, somatosensory feedback control typically contributes little to the speech of fluent adult speakers, since under normal settings feedforward motor programs for frequently produced syllables are highly tuned through practice and thus speakers make few, if any, significant somatosensory errors. However, an unexpected somatosensory perturbation will lead to a difference between the speaker's somatosensory expectations of the current speech sound and the actual, afferent feedback (Golfinopoulos et al., 2010; Guenther, 1994, 1995, 2006; Guenther et al., 2006; Tourville and Guenther, 2010). The DIVA model predicts that under these conditions in which somatosensory feedback is not within the expected range for the current speech sound there is increased activation in bilateral supramarginal gyrus (the location of the model's *somatosensory error map*). In earlier versions of the DIVA model, somatosensory errors were transformed into compensatory motor commands via projections to bilateral ventral motor cortex (e.g., Guenther, 2006; Guenther et al., 2006). A recent fMRI study of auditory feedback control of speech demonstrated the involvement of right prefrontal and ventral premotor areas in generating corrective speech motor commands (Tourville et al., 2008). Increased bilateral activity in posterior superior temporal gyrus and ventral motor cortex was accompanied by increases in right ventral premotor cortex and inferior frontal gyrus, *pars triangularis* when auditory feedback was unpredictably perturbed during speech production relative to speech. Structural equation modeling of this network of regions revealed significantly greater effective connectivity between bilateral posterior superior temporal gyrus and right inferior frontal gyrus *pars triangularis* and right ventral premotor cortex for perturbed speech relative to unperturbed speech. These findings suggest that error signals detected in bilateral sensory cortex were transformed into compensatory speech motor commands via projections through right inferior frontal and ventral premotor cortex. This interpretation is reflected in the most recent version of the DIVA model (e.g., Golfinopoulos et al., 2010; Tourville and Guenther, 2010). According to the model, an unanticipated *somatosensory* perturbation will result in increased bilateral activity in anterior supramarginal gyrus associated with somatosensory error detection and right-lateralized ventral premotor cortex and inferior frontal gyrus *pars triangularis* along with bilateral ventral motor cortex activity associated with the compensatory motor response. Furthermore, the effective connectivity between anterior supramarginal gyrus and ventral premotor cortex and inferior frontal gyrus *pars triangularis* should increase as a result of the perturbation.

To test these predictions, we measured blood oxygen level dependent (BOLD) responses during *speech* and *baseline* tasks under unperturbed and perturbed conditions using event-related sparse sampling fMRI. Subjects produced bi-syllabic pseudowords aloud in periods of silence inside the scanner. During a subset of the *speech* and *baseline* trials, a pneumatic device unpredictably blocked subjects' jaw closing movements. These unanticipated perturbations prompted subjects to adjust articulator trajectories in order to continue to produce the visually presented speech stimulus. BOLD activity associated with this perturbation during speech (*perturbed speech*) was contrasted with that under unperturbed speech conditions (*speech*) using both voxel-based and region of interest-based functional analyses. In addition, structural equation modeling was used to quantitatively assess changes in the effective connectivity between the *perturbed speech* and *speech* conditions. Path or structural equation modeling analysis has been used extensively on neuroimaging data to test specific hypotheses about regional connectivity (e.g., Au Duong et al., 2005; Gonçalves et al. 2001; Grafton et al., 1994; Horwitz et al., 1995; Rowe et al., 2002; Schlösser et al., 2003; Tourville et al., 2008). However, SEM is sensitive to the choice of regions and the complexity of the structural model (e.g., limitations on the number of paths that can be fitted; for further discussion see Penny et al., 2004). To mitigate these limitations, SEM is used here as a confirmatory tool; the network modeled and hypotheses tested are constrained by our prior theoretical and experimental work.

Materials and methods

Subjects

Participants were 13 right-handed native speakers of American English (6 females; age range 23–51, mean age = 30) with corrected or normal vision. None of the subjects reported a neurological, speech, hearing, or voice impediment. All subjects were recruited and provided written informed consent in accordance with the guidelines of the Boston University Institutional Review Board and the Massachusetts General Hospital Human Research Committee. Subjects received monetary compensation for their participation in this study.

Experimental protocol

During functional magnetic resonance imaging, subjects performed an overt speech production task and an observational baseline

task. In the *speech* task, subjects were instructed to read aloud the stimulus (a pseudoword) that was orthographically presented for three seconds at the beginning of each trial. The speech materials consisted of two /aV/ pseudowords (*au, ai*) and six /aCV/ pseudowords (*anu, ani, agu, agi, atu,* and *ati*). During the *baseline* task, a control stimulus (*yyy*) was presented in place of the pseudoword, and subjects were instructed to silently observe the stimulus throughout the three second presentation. Subjects could view all stimuli, which were projected onto a display outside the MR scanner bore, via an adjustable mirror positioned on the head coil. The stimuli were presented in white against a black background in the center of the display. All stimuli were non-lexical so as to avoid possible confounds from semantic effects.

During stimulus presentation the scanner remained silent and a Shure SM93 Micro-Lavelier electrostatic microphone attached to the head coil (approximately 3 in. from the subject's mouth) transmitted subjects' speech productions to a computer inside the control room. On a subset of all *speech* and *baseline* trials, denoted *perturbed speech* and *perturbed non-speech* trials, jaw movements were blocked by the rapid inflation of a small balloon (for a description of the perturbation setup, see Somatosensory perturbation section). Presentation software package (version 0.80; www.neurobs.com) was used to deliver the visual stimuli and to trigger both the perturbation and the scanner for image acquisition. Each experimental run consisted of 56 *speech* trials (seven presentations of each pseudoword) and 16 *baseline* trials. Subjects were asked to perform four, 13-minute runs in a single scanning session. For two subjects, only three runs were used in the analysis due to technical difficulties at the time of image acquisition. Subjects wore protective ear covering throughout the experiment.

Somatosensory perturbation

Throughout the experiment a small, approximately tubular, inelastic balloon (Fig. 1B) was positioned between the molars on one side of each subject's mouth. For six of the subjects the balloon was positioned on the left side of the mouth (hereafter called the *left balloon cohort*), and for the remaining seven subjects the balloon was positioned on the right side of the mouth (the *right balloon cohort*). The balloon was constructed from a finger of a heavy-duty nitrile rubber glove. Prior to scanning, subjects evaluated the fit of various balloons and chose the most comfortable shape for their mouth. None of the subjects reported that the uninflated balloon prevented speech-related movements. The balloon was connected to a long stiff plastic tube attached to a solenoid-driven air cylinder that was located inside the scanner control room to avoid providing auditory cues to the subject regarding the onset of the jaw perturbation (Fig. 1A). Computer-triggered activation of the solenoid delivered 4–5 psi and caused inflation of the balloon to a thickness of approximately 1 cm within 100 ms.

The balloon was inflated on one half of all *baseline* trials and one seventh of all *speech* trials pseudo-randomly distributed within each run so that the presentation of two consecutive perturbation trials was prohibited. The frequency of perturbation trials was chosen to avoid adaptation to, or anticipation of, the perturbation on any given trial. Previous dynamic perturbation experiments indicate that using 15% or fewer perturbations will prevent subjects from performing anticipatory motor acts that minimize the influence of the perturbation (Abbs and Gracco, 1984; Gracco and Abbs, 1985). Before the start of the experiment, subjects were informed of the possibility that the balloon would inflate during some trials, but that they should continue to produce the stimulus and avoid clenching their teeth. Subjects were instructed to keep their jaws relaxed and in a closed position during *baseline* trials and between trials.

Production of each pseudoword required articulator movements that are predictably disrupted by a jaw block. Specifically, all stimuli started with the vowel /a/, which under normal conditions involves a low (open) jaw position, and transitioned to consonants or vowels that involve high (relatively closed) jaw positions. During *perturbed speech* trials, the balloon inflation was triggered at the onset of voicing at the start of the /a/ (while the jaw was open, so the balloon expansion had very little effect) and continued through the remainder of the utterance, thereby impeding the closing jaw movement for the subsequent phoneme(s). Before entering the scanner, subjects were instructed to speak each word slowly and clearly and practiced the pronunciation of the stimuli until the subject's production consistently matched a sample production. Inside the scanner, subjects' productions were monitored to ensure proper performance and timing of the perturbation. For *baseline* trials, where voicing was not available as a cue, triggering of the perturbation was initiated 500 ms after the onset of the trial and remained inflated until the baseline stimulus disappeared from the screen.

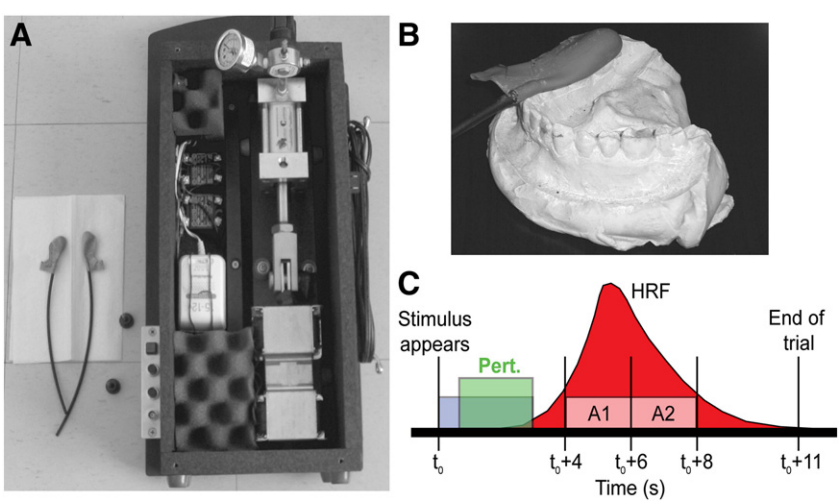

Fig. 1. A. The perturbation apparatus. A solenoid-driven air cylinder was used to apply an unanticipated mechanical load to the subject's jaw. A tubular shaped balloon, pictured on the *left*, was custom fashioned to fit comfortably between the molar teeth on one side of the mouth as demonstrated in B. On perturbed trials, the air cylinder, located in the scanner control room, delivered 4–5 psi to the balloon via a long stiff plastic tube, causing the balloon to inflate to a diameter of about 1 cm within 100 ms. The inflated balloon blocked closure movements of the jaw. C. Timeline of a single trial in the event-triggered sparse sampling protocol. At the onset of each trial, the visual stimulus appeared and remained onscreen for 3 s (blue rectangle). On perturbed trials, the balloon inflation was triggered after the onset of the trial and remained inflated until the stimulus disappeared from the screen (green). About 1 s after stimulus offset, two whole-brain volumes were acquired (A1 and A2). Data acquisition was timed to cover the peak of the hemodynamic response to speech; the putative hemodynamic response function (HRF) is schematized in red. The next trial started 3 s after data acquisition was complete, resulting in a total trial length of 11 s.

Behavioral pilot experiment

In order to ensure that subjects could compensate for the restrained jaw induced by the pneumatic device, a separate behavioral pilot study was performed on a healthy, native speaker of American English (1 male) prior to the fMRI experiment. The subject was recruited and provided written informed consent in accordance with the guidelines of the Massachusetts Institute of Technology Review Board and received monetary compensation for his participation. The subject was comfortably seated in a modified dental chair in a quiet recording room and transducers were mounted on midsagittal points of the tongue body, upper and lower lips, and mandible. Kinematic data were collected using an electro-magnetic midsagittal articulometer that measured the positions of these transducers during *speech* and *perturbed speech* conditions. After becoming familiar with the perturbation setup and the stimuli, the subject performed a total of 248 *speech* trials (31 unperturbed productions of each speech stimulus) and 32 *perturbed speech* trials (4 productions of each stimulus in which jaw movement was unexpectedly perturbed) within a single run.

MRI data acquisition

Magnetic resonance imaging was performed on a Siemens Trio 3T whole-body scanner with a volume transmit–receive birdcage head coil (USA Instruments, Aurora, OH) at the Massachusetts General Hospital Athinoula A. Martinos Center for Biomedical Imaging in Charlestown, MA. Foam padding was packed around the subject's head to minimize head movement during the experiment. For each subject, a high resolution T1-weighted anatomical volume (128 slices in the sagittal plane, slice thickness = 1.33 mm, in-plane resolution = 1 mm^2, TR = 2000 ms, TE = 3300 ms, flip angle = 7°, FOV = 256 mm^2) was obtained prior to functional imaging. Functional volumes consisted of 32 T2*-weighted gradient echo, echo planar images covering the whole brain in the axial plane, oriented along the bicommissural line (slice thickness = 5 mm, in plane resolution = 3.125 mm^2, no skip, TR = 2000 ms, TE = 30 ms, flip angle = 90°, FOV = 200 mm^2).

A sparse sampling (Hall et al., 1999), clustered volume acquisition method, consisting of silent intervals between consecutive volume acquisitions, was used. Subjects performed the tasks during the silent intervals. The timeline for a single trial is schematized in Fig. 1C. Functional acquisition began 4 s after the onset of the trial, following the completion of the task and consisted of two consecutive whole-brain functional volumes (TR = 2 s for each volume) within each trial. Acquisition timing was chosen to include the putative peak in the delayed hemodynamic response associated with task performance; the hemodynamic response delay has been estimated at 4–7 s, depending on the brain region and task (Belin et al., 1999; Yang et al., 2000). After the two volumes were collected, there was a 3 s interval before the onset of the following trial to allow for attenuation of the hemodynamic response to scanner noise between trials, resulting in a total trial length of 11 s. The sparse sampling protocol permitted speech production in the absence of high frequency, high intensity acquisition-related acoustic noise (Belin et al., 1999). It also minimized movement-related imaging artifacts, since scanning occurred only after the behavioral task had ended (Birn et al., 1999).

MRI data analyses

Voxel-based analysis

Functional imaging data were processed using tools from the SPM2 software package (Wellcome Department of Imaging Neuroscience, University College of London; http://www.fil.ion.ucl.ac.uk/spm/). Functional images were realigned to the mean echo-planar image (EPI) within a session by estimating the parameters of an optimal rigid body transformation. Functional images were then co-registered to the subject's T1-weighted anatomical image. Each subject's anatomical image was spatially normalized to the Montreal Neurological Institute T1 template (MNI ICBM-152; Evans et al., 1993). The same spatial normalization transformation applied to the anatomical image was also applied to the functional images. Global scaling was used to eliminate the effects of low frequency noise sources. Functional images were spatially smoothed using an isotropic Gaussian kernel of 8 mm full-width at half maximum.

Task-dependent BOLD responses were estimated using a general linear model. The hemodynamic response for each stimulus event was modeled using a finite impulse response (FIR) basis function with a single time bin spanning the two consecutive volumes. The model included four condition-specific explanatory variables (*speech*, *perturbed speech, perturbed non-speech,* and *baseline)*. Also included in the model were motion parameters estimated during the realignment step, a high-frequency regressor ($[1\ -1\ 1\ -1\ \ldots]^T$) modeling global intensity differences between the first and second volumes collected each trial, and a linear regressor to account for scanner drift.

The model was first estimated within each subject, and contrast-of-interest volumes were generated comparing the appropriate conditional coefficients to assess the four simple main effects and the interaction term with task (speech versus non-speech production) and modulation (perturbation versus no perturbation) as the two factors. Specifically, inferences were based on the five contrasts: *speech–baseline*, *perturbed non-speech–baseline*, *perturbed speech–perturbed non-speech*, *perturbed speech–speech,* and (*perturbed speech–speech*)–(*perturbed non-speech–baseline*). To assess differences in the effects of balloon position, a second level, two-sample *t*-test was performed to compare the *left* and *right balloon cohorts* for the five main contrasts of interest. A false discovery rate (FDR; Genovese et al., 2002) corrected threshold of $p_{FDR}<5\%$ was used to asses significance.

Direct comparisons between condition-specific brain activations were then performed using a second level, mixed effects analysis across all subjects (collapsing *left* and *right balloon cohorts*). The resulting group statistical parametric maps were thresholded by a corrected significance of $p_{FDR}<1\%$.

For visualization/reporting purposes, those voxels surpassing the significant *t* threshold, effect sizes were divided by the mean significant effect ($p<0.01$, uncorrected) of the *perturbed speech–speech* contrast (the primary contrast of interest) to generate *normalized effect sizes*. This normalized effect size demonstrates relative activations between contrasts for voxels that survive a significance threshold. The normalized effect sizes were plotted on a cortical surface representation of the canonical anatomical volume included in the SPM2 package to visualize significant effects for each contrast of interest.

In addition to generating statistical parametric maps of the voxel-based results, peak responses were associated with neuroanatomical labels by mapping them to the same set of cortical, subcortical, and cerebellar regions used in region-of-interest (ROI)-based functional analyses described below. Peaks in the *t*-statistic maps (peak-to-peak minimum distance = 6 mm) were assigned ROI labels based on a minimum distance function. The function first determines, in a subject-specific manner, the nearest (in Euclidean distance) ROI in the individual subjects' space for a given peak voxel location in the common coordinate space. Since the nearest region can differ across subjects, this procedure results, for each peak, in a set of candidate nearest regions with size less than or equal to the number of subjects. Then, for each of these candidate regions, we determine the mean Euclidean distance between the peak location and the nearest voxel within that region across all subjects. The candidate region with the minimum average distance across subjects is then associated with the given peak voxel. For each ROI containing a peak response, the voxel location, *t*-statistic, and normalized effect of the maximum response are determined. MNI coordinates were converted into Talairach

coordinates using an MNI to Talairach transformation detailed in Lancaster et al. (2007).

Region of interest based analysis

ROI analyses were used to test neuroanatomically-specific *a priori* hypotheses regarding activity in the speech motor control network under conditions in which somatosensory feedback is unexpectedly perturbed during speech production. For the present study, BOLD responses were assessed within a set of 39 pre-defined cortical, subcortical, and cerebellar ROIs[1] in each hemisphere. A schematic of the ROIs is shown in Supplementary Fig. 1. Tests performed were used to determine which of these regions showed significant effects for each contrast of interest, to test the set of regions for significant laterality effects within each contrast of interest, and to extract data for effective connectivity analysis.

The ROI analysis eliminated the need for the non-linear spatial normalization and smoothing steps of the voxel-based analysis; image pre-processing was otherwise identical to that of the voxel-based analysis. Freesurfer image processing software (http://surfer.nmr.mgh.harvard.edu) was used to segment gray and white matter structures (Fischl et al., 2002) and to reconstruct cortical surfaces from each subject's anatomical volume (Dale et al., 1999; Fischl et al., 1999). Subcortical ROIs were parcellated according to a training set provided by Freesurfer (Fischl et al., 2002). Cortical and cerebellar regions were parcellated based on a parcellation scheme tailored specifically for speech studies (Tourville and Guenther, 2003) using Freesurfer (Fischl et al., 2004). ROI masks were inspected following Freesurfer classification by an experienced human rater (JAT), and manual edits were made as needed.

Analysis of the BOLD responses within each ROI was performed according to the procedure described by Nieto-Castanon et al. (2003). Briefly, BOLD responses were averaged across all voxels within each ROI mask, and these mean regional responses for each stimulus event were modeled using a finite impulse response (FIR) basis function with a single time bin spanning the two consecutive volumes. The same set of condition-related regressors used in the voxel-based analyses (*speech*, *perturbed speech, perturbed non-speech,* and *baseline)* were used in the design matrix of the ROI analyses to model the average regional responses.

Group level effects were assessed by first computing regional contrasts for each subject. The regional contrasts were then pooled across subjects and individual ROIs were tested for significance using one-sample *t*-tests and thresholded at $p_{FDR}<5\%$. Since a voxel-wise FDR correction can typically be more liberal than a region-wise FDR correction (cf., Chumbley and Friston, 2009) we applied a less conservative threshold of $p<5\%$ corrected for multiple comparisons for the ROI analyses. Paired left and right ROIs were subsequently tested for laterality effects using a paired *t*-test. In a first set of laterality tests, *a priori* hypotheses that responses in inferior frontal gyrus *pars opercularis*, inferior frontal gyrus *pars triangularis* and ventral premotor cortex are significantly left-lateralized under normal speech conditions (cf. Ghosh et al., 2008; Tourville et al., 2008) were tested. Based on the findings from the *auditory* feedback control study, we also tested the hypothesis that responses in the ventral premotor cortex would become right-lateralized when compensatory responses are induced (cf. Tourville et al., 2008). Significance for the three laterality tests was determined using a threshold of $p<5\%$. Laterality effects for the remaining set of regions were tested using paired *t*-tests and thresholded at $p_{FDR}<5\%$. Regional effect sizes for all ROIs were normalized by the mean significant ($p<5\%$) effect of the *perturbed speech–speech* contrast and visualized on a bar plot.

Structural equation modeling

Structural equation modeling (SEM) was performed to assess inter-regional effective connectivity between five cortical regions hypothesized to be part of the somatosensory feedback control network for speech production (see Golfinopoulos et al. (2010) including bilateral anterior supramarginal gyrus, right inferior frontal gyrus, *pars triangularis*, ventral premotor cortex, and ventral motor cortex; anatomical boundaries for these ROIs are provided in Supplementary Table 1 and highlighted in gray in Supplementary Fig. 1). ROI time series data were generated by first determining the regional response for each of the five ROIs in each functional volume. Each time series was limited to those voxels with an effect size in the top 50th percentile for that region in the *speech–baseline* contrast (for the number of voxels selected for each region for each subject, see Supplementary Table 2). Regional responses from each functional run were divided into two series, one consisting of the first volume collected in each trial and the other consisting of the second volume collected. The two series were each linearly detrended and then averaged, yielding a single regional response for each trial, subject, and region of interest. These values were grouped by condition within each subject and then concatenated to form a time series containing average trial responses for each condition and region. Outliers (greater than three standard deviations) were removed, and the data were standardized with zero mean and unit variance.

Tabulated ROI data for the *speech* and *perturbed speech* conditions were used as input for the structural equation modeling analysis performed with AMOS 7 software (http://www.spss.com/amos/index.htm). SEM estimates interregional effective connectivity by minimizing the difference between observed regional *covariances* and those implied by a structural model (cf., Büchel and Friston, 1997). It was applied here to test whether effective connectivity between bilateral anterior supramarginal gyrus and right ventral premotor cortex and inferior frontal gyrus in the hypothesized somatosensory feedback control network (shown in Fig. 4) increased when somatosensory feedback was perturbed during speech. Free parameters of the structural model (e.g., path coefficients and residual effects) were estimated by minimizing a maximum likelihood (ML) discrepancy function. To assess the overall fit of the model to the observed data, a chi-square test was performed since the minimum of the ML discrepancy function multiplied by the number of observations minus one is chi-square distributed with $(q/2)(q+1)$-p degrees of freedom where (p is the number of estimated parameters and q is the number of observed variables; Büchel and Friston, 1997). The objective of this chi-square test is to *confirm* the null hypothesis that the model is correct in the population (Schlösser et al., 2006). However, since the χ^2 test is sensitive to sample size and the distribution of the data, four additional goodness-of-fit indices were also used in the present study: the goodness-of-fit index (GFI), the adjusted goodness-of-fit (AGFI), the root mean square residual (RMR), and the root mean square error of approximation (RMSEA).

In order to test for *task-specific* changes in effective connectivity we used a stacked model approach (Büchel and Friston, 1997) comparing a *null model* (in which all free parameters are constrained to be equal across the two conditions) and an *alternative model* (in which path coefficients are permitted to vary between the two conditions). The residual variances were constrained to be equal across the conditions in both the null and alternative models to reduce

[1] The following set of speech-related cortical and subcortical ROIs was utilized based on a review of neurophysiological and imaging studies of speech processing (Tourville and Guenther, 2003): inferior frontal gyrus *pars opercularis*, inferior frontal gyrus *pars triangularis,* frontal operculum, ventral premotor cortex, anterior and posterior central operculum, ventral motor cortex, ventral somatosensory cortex, anterior and posterior supramarginal gyrus, parietal operculum, anterior cingulate gyrus, pre-supplementary motor area, supplementary motor area, anterior and posterior insula, Heschl's gyrus, planum polare, planum temporale, anterior and posterior superior temporal gyrus, anterior and posterior middle temporal gyrus, middle temporal occipital gyrus, anterior dorsal, anterior ventral, posterior dorsal, and posterior ventral superior temporal sulcus, anterior medial cerebellum, anterior lateral cerebellum, superior posterior medial cerebellum, superior posterior lateral cerebellum, inferior posterior medial cerebellum, inferior posterior lateral cerebellum, thalamus, caudate, putamen, and pallidum.

the number of estimated parameters. A χ^2 difference test was performed in which the χ^2-value for the alternative model was subtracted from the χ^2-value of the null model to determine whether the alternative model provides a significantly better fit to the data than does the null model. A significant χ^2_{diff} -value ($p<5\%$) provides evidence that the null model, in which parameters are constrained to be the same for the two conditions, should be rejected, thus providing evidence that the effective connectivity significantly differs across certain paths between the two conditions. In order to determine which path coefficients were statistically significantly different between the two conditions pair-wise comparisons of path coefficients for the two conditions were performed using two-tailed *z*-tests.

Results

Behavioral responses

Although it is not currently possible to measure kinematic data for the articulators inside the MR scanner, kinematic data of a single subject were recorded in a pilot experiment prior to the fMRI experiment to determine whether subjects could compensate for the dynamic jaw perturbation. The results of this pilot indicated that the subject's average jaw height (averaged across all trials of a particular stimulus) for each stimulus was 1.3–2.4 mm lower during the final vowel on *perturbed speech* trials than on *speech* trials, while the average tongue blade height during the final vowel was nearly identical in the *perturbed speech* and *speech* conditions (average differences of less than 0.2 mm for all stimuli). These data suggest that the subject compensated for the downward jaw perturbation by moving the tongue upward relative to the jaw in order to achieve the oral cavity constriction necessary for appropriate acoustic production of the final vowel. Increased upward movements of the tongue to accomplish oral constriction in response to unexpected jaw perturbations during speech is consistent with earlier reports of behavioral studies that indicate the ability of speakers to successfully compensate for sudden jaw perturbations (e.g., Folkins and Abbs, 1975; Gomi et al., 2002; Ito et al., 2005; Kelso et al., 1984; Shaiman, 1989).

Neural responses

Effects of balloon placement

Before performing direct comparisons between condition-specific brain activations, we assessed differences in the neural responses of the two cohorts of subjects with the balloon positioned on the right and left side of the mouth. No voxels survived the two-sample *t*-tests ($p_{FDR}<5\%$) for differences between the left and right balloon cohorts ($n_{left\ balloon\ cohort}=6$, $n_{right\ balloon\ cohort}=7$) in any of the contrasts of interest (*speech–baseline*, *perturbed non-speech–baseline*, *perturbed speech–perturbed non-speech, perturbed speech–speech,* (*perturbed speech–speech*)–(*perturbed non-speech–baseline*)), suggesting that the average brain activation patterns did not significantly depend on the placement of the balloon within the speakers' mouths. The condition effects described below were based on the group data from both cohorts.

Effects of condition

Fig. 2A presents the normalized group effects of the *speech–baseline* contrast on a cortical surface rendering following a mixed

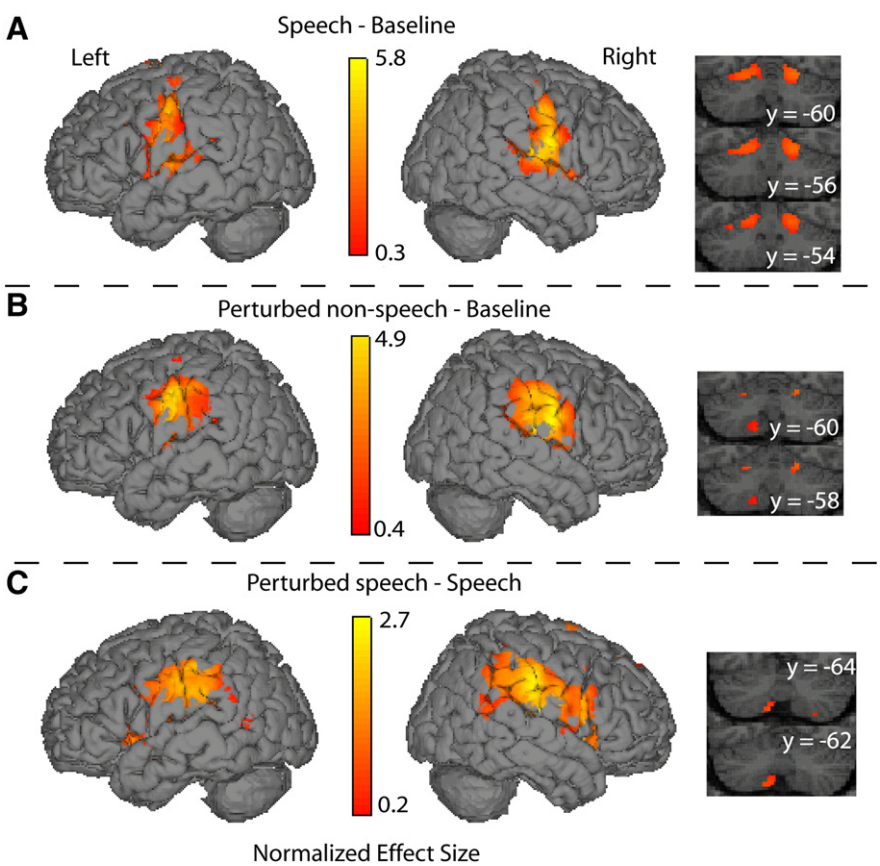

Fig. 2. Group maps displaying the normalized effect sizes of voxels that surpassed the threshold ($t_{speech-baseline,\ 12}>4.13$; $t_{perturbed\ non-speech-baseline,\ 12}>4.34$; $t_{perturbed\ speech-speech,\ 12}>4.32$; $p_{FDR}<1\%$) for (A) *speech* relative to *baseline*, (B) *perturbed non-speech* relative to *baseline*, and (C) *perturbed speech* relative to *speech*. BOLD responses are overlaid on the SPM2 single-subject canonical T1 dataset. Coronal slices through the cerebellum are shown to the right of the lateral cortical surfaces. The anatomical *left* is on the *left* side of each coronal image; the *y* coordinate indicates the anterior–posterior location of each slice in MNI space. In the cortical renderings of the *speech–baseline*, *perturbed non-speech–baseline*, and *perturbed speech–speech* contrasts, the highest normalized effect shifts from the left to the right hemisphere and, then, posteriorly into anterior supramarginal gyrus.

Table 1
Peak voxel responses for the three contrasts listed by anatomical region.

Region label		*Speech–Baseline*				*Perturbed non-speech–Baseline*				*Perturbed speech–Speech*					
		Peak voxel location (*x*, *y*, *z*)		T	Normalized effect	Peak voxel location		T	Normalized effect	Peak voxel location		T	Normalized effect	ROI T	ROI normalized effect
		MNI	Talairach			MNI	Talairach			MNI	Talairach				
Rolandic cortex															
Left	pdPMC	(−6, −12, 74)	(−7, −19, 69)	6.49	1.05					(−50, −6, 38)	(−48, −10, 37)	8.13	0.52	n.s.	n.s.
	vPMC									(−60, 8, 30)	(−57, 3, 31)	5.99	1.16	4.28	0.87
	dMC	(−56, −12, 50)	(−53, −17, 47)	9.05	3.46	(−44, −14, 62)	(−43, −20, 58)	5.42	0.94	(−56, −8, 38)	(−53, −12, 36)	7.24	1.31	n.s.	n.s.
	vMC	(−62, −4, 34)	(−59, −8, 33)	10.65	2.89	(−62, −2, 36)	(−59, −6, 35)	9.83	2.74	(−58, −6, 30)	(−55, −10, 29)	6.45	1.55	5.40	1.63
	vSC	(−58, −10, 40)	(−55, −14, 38)	10.66	4.29	(−58, −10, 42)	(−55, −14, 40)	11.61	3.89	(−58, −18, 24)	(−55, −20, 23)	8.70	1.81	7.40	1.93
	pCO	(−56, −10, 10)	(−53, −12, 11)	8.05	2.43	(−46, −10, 22)	(−44, −13, 22)	6.27	1.53	(−52, −12, 12)	(−49, −14, 13)	6.68	1.09	4.88	1.45
Right	mdPMC									(18, −2, 66)	(15, −9, 63)	6.62	0.86	3.95	0.75
	pdPMC	(62, −6, 46)	(56, −12, 46)	6.87	2.92	(8, −8, 72)	(6, −16, 68)	5.64	1.11						
	vPMC									(62, 16, 32)	(56, 10, 35)	5.74	1.08	4.63	1.51
	dMC	(46, −12, 42)	(41, −17, 41)	6.56	3.01					(54, −8, 36)	(49, −13, 36)	6.61	1.03	n.s.	n.s.
	vMC	(60, −6, 36)	(54, −11, 37)	7.36	4.00	(64, 6, 28)	(58, 1, 31)	8.90	2.13	(66, 10, 14)	(60, 6, 18)	7.88	1.40	5.16	1.61
	vSC	(66, −4, 30)	(60, −8, 32)	6.73	3.97	(62, −8, 36)	(56, −13, 36)	12.06	4.02	(54, −20, 42)	(48, −24, 41)	7.03	1.26	6.18	1.76
	dSC	(52, −12, 58)	(46, −18, 56)	4.54	1.78										
	aCO	(44, 6, 6)	(40, 3, 11)	7.31	1.30	(44, −4, 12)	(40, −7, 15)	7.73	2.82	(46, 0, 18)	(41, −4, 21)	7.37	1.64	5.50	1.85
	pCO	(48, −12, 22)	(43, −15, 23)	10.94	1.25	(48, −10, 18)	(43, −13, 20)	7.84	2.16	(46, −10, 16)	(41, −13, 18)	6.21	1.59	4.48	1.71
Frontal cortex															
Left	preSMA	(−4, 6, 46)	(−5, 0, 46)	6.41	1.72					(−4, 8, 56)	(−5, 1, 55)	5.82	0.57	n.s.	n.s.
	SMA	(−6, −2, 46)	(−7, −8, 45)	6.59	1.42					(−6, 0, 54)	(−7, −6, 52)	4.69	0.82	n.s.	n.s.
	aCg	(−2, 24, 28)	(−3, 18, 31)	6.90	0.62										
	pCg									(−10, −24, 44)	(−11, −28, 41)	4.99	0.79	n.s.	n.s.
	adPMC	(−8, 4, 76)	(−9, −5, 72)	5.10	1.85										
	IFo									(−52, 12, 2)	(−49, 9, 6)	6.69	1.24	3.35	0.60
	FO									(−40, 20, 12)	(−38, 16, 16)	5.00	0.63	4.13	1.12
	FOC									(−32, 18, −14)	(−30, 16, −7)	4.86	0.84	n.s.	n.s.
Right	pMFg									(52, 14, 46)	(47, 7, 47)	4.46	0.68	n.s.	n.s.
	SFg									(6, 30, 58)	(4, 21, 59)	7.20	1.13	3.65	0.67
	preSMA	(10, 2, 48)	(8, −4, 47)	4.40	0.94	(4, 0, 52)	(2, −6, 51)	6.27	2.32	(2, 10, 58)	(0, 2, 57)	7.46	1.03	4.83	1.05
	SMA	(0, −2, 68)	(−2, −10, 65)	6.45	3.24	(2, −4, 60)	(0, −11, 57)	5.62	2.51						
	aCg	(2, 20, 36)	(0, 14, 38)	8.73	1.10	(2, 8, 40)	(0, 2, 41)	5.28	1.12	(6, 2, 40)	(4, −3, 40)	6.11	1.20	4.07	0.68
	IFo									(50, 20, 6)	(45, 16, 12)	5.78	1.29	4.16	1.22
	IFt									(50, 20, 0)	(45, 17, 7)	6.05	1.54	4.74	0.93
	FOC									(42, 20, −12)	(38, 18, −4)	8.75	1.16	n.s.	n.s.
	FP									(12, 50, 48)	(10, 41, 52)	5.10	0.76	n.s.	n.s.
Parietal cortex															
Left	aSMg					(−64, −22, 30)	(−61, −25, 28)	13.21	2.36	(−66, −30, 30)	(−63, −32, 27)	8.60	1.87	7.07	1.65
	pSMg									(−64, −52, 30)	(−61, −53, 25)	4.65	0.56	5.45	0.77
	PO	(−46, −38, 26)	(−44, −39, 23)	8.24	1.64	(−46, −36, 24)	(−44, −37, 22)	7.77	2.29	(−54, −32, 22)	(−51, −33, 20)	5.09	1.39	5.43	1.77
	AG									(−64, −52, 24)	(−61, −52, 20)	4.92	0.60	n.s.	n.s.
Right	aSMg					(60, −22, 26)	(54, −25, 26)	13.41	3.66	(62, −20, 26)	(56, −23, 26)	7.49	2.36	6.00	2.33
	pSMg					(56, −32, 28)	(50, −34, 27)	14.31	2.71	(68, −38, 34)	(61, −40, 32)	11.29	1.44	5.02	1.63
	PO	(48, −30, 22)	(43, −32, 22)	6.88	2.00										
	SPL									(40, −42, 58)	(35, −46, 53)	6.33	0.66	n.s.	n.s.
	AG									(60, −52, 18)	(54, −52, 16)	7.10	1.04	4.25	0.89

Temporal cortex															
Left	Hg	(−48, −20, 10)	(−46, −21, 11)	7.38	1.91	(−42, −18, 6)	(−40, −19, 7)	6.51	1.13	(−56, −18, 10)	(−53, −19, 11)	5.78	0.93	n.s.	n.s.
	aSTg									(−60, 0, −12)	(−56, 0, −8)	4.96	0.48	n.s.	n.s.
	pSTg	(−68, −32, 10)	(−64, −32, 9)	4.56	1.52										
	PT					(−56, −42, 22)	(−53, −43, 19)	7.55	0.85	(−60, −38, 20)	(−57, −39, 18)	4.86	1.41	3.33	0.51
	PP									(−60, 4, −6)	(−56, 3, −2)	4.97	0.60	n.s.	n.s.
	MTO									(−62, −58, 10)	(−59, −56, 7)	5.34	0.62	4.38	0.59
Right	TP	(56, 8, −6)	(51, 6, 0)	7.00	1.46					(56, 14, −12)	(51, 12, −5)	5.63	0.71	n.s.	n.s.
	Hg	(48, −20, 10)	(43, −22, 12)	9.06	2.28	(42, −24, 14)	(38, −26, 15)	5.40	1.39						
	pSTg	(68, −22, 2)	(62, −23, 5)	5.73	1.54										
	PT	(60, −14, 4)	(54, −15, 7)	8.77	2.02										
	PP					(44, −10, −6)	(40, −11, −1)	8.24	1.43	(42, −26, 4)	(38, −27, 6)	5.55	0.30	3.69	1.06
	MTO					(58, −58, −2)	(53, −56, −2)	4.49	0.91	(54, −58, 8)	(49, −57, 7)	4.97	0.92	3.91	0.98
Insular cortex															
Left	aINS	(−40, −6, 8)	(−38, −8, 10)	13.70	1.02	(−38, −4, 14)	(−36, −7, 16)	11.04	2.31	(−40, −2, 0)	(−38, −3, 3)	12.05	1.73	6.23	1.43
	pINS									(−40, −20, 8)	(−38, −21, 9)	5.11	0.89	4.50	1.00
Right	aINS	(36, 4, 4)	(32, 1, 9)	7.51	0.89	(46, −2, −2)	(42, −4, 3)	7.63	2.39	(44, −2, 4)	(40, −4, 8)	6.84	2.19	4.86	1.30
	pINS									(42, −6, −6)	(38, −7, −1)	5.61	1.45	4.17	0.94
Cerebellum															
Left	amCB, L5	(−12, −54, −18)	(−12, −50, −17)	5.41	1.71										
	splCB, L6	(−26, −60, −22)	(−25, −52, −20)	5.20	1.99	(−22, −58, −24)	(−21, −54, −23)	5.12	1.36						
	ipmCB, L8					(−16, −60, −52)	(−15, −53, −48)	5.36	0.91	(−8, −62, −48)	(−8, −58, −45)	6.10	0.76	n.s.	n.s.
Right	spmCB, L6	(18, −56, −18)	(16, −52, −17)	6.20	2.44										
	splCB, L6					(20, −58, −24)	(18, −54, −22)	4.74	1.77						
	iplCB, L8									(28, −64, −56)	(25, −57, −51)	4.52	0.48	n.s.	n.s.
Subcortical nuclei															
Left	Pal	(−22, 0, −6)	(−21, −1, −2)	8.02	1.49										
	Put	(−24, −6, 8)	(−23, −8, 10)	9.68	1.21	(−24, 2, −4)	(−23, 1, 0)	4.55	1.18						
	Tha, MD	(−12, −18, 8)	(−12, −19, 9)	7.24	1.50					(−2, −22, 4)	(−3, −23, 6)	5.07	1.11	n.s.	n.s.
	Tha, VPM	(−14, −20, −2)	(−14, −20, 0)	6.32	1.05										
Right	Pal	(18, 0, 4)	(16, −2, 8)	8.97	0.62										
	Put	(30, −12, 6)	(27, −14, 9)	9.40	0.86	(24, 0, −10)	(21, −1, −4)	5.17	1.81						
	Caud	(24, −12, 20)	(21, −15, 21)	5.06	0.51										
	Tha, MD	(10, −12, 10)	(8, −14, 12)	6.97	1.33	(12, −12, 10)	(10, −14, 12)	5.32	1.25	(10, −18, 12)	(8, −20, 13)	5.23	0.84	n.s.	n.s.
	Tha, VPM	(18, −22, −2)	(16, −22, 1)	6.11	0.99					(16, −22, 0)	(14, −22, 2)	5.04	0.60	n.s.	n.s.

Abbreviations: aCg = anterior cingulate gyrus; aCO = anterior central operculum; adPMC = anterior dorsal premotor cortex; AG = angular gyrus; aINS = anterior insula; amCB = anterior medial cerebellum; aSMg = anterior supramarginal gyrus; aSTg = anterior superior temporal gyrus; Caud = caudate; dMC = dorsal primary motor cortex; dSC = dorsal somatosensory cortex; FO = frontal operculum; FOC = frontal orbital cortex; FP = frontal pole; Hg = Heschl's gyrus; IFo = inferior frontal gyrus, pars opercularis; IFt = inferior frontal gyrus, pars triangularis; iplCB = inferior posterior lateral cerebellum; L5 = cerebellum lobule V; L6 = cerebellum lobule VI; L8 = cerebellum lobule VIII; MD = mediodorsal thalamic nucleus; mdPMC = middle dorsal premotor cortex; MTO = middle temporal occipital gyrus; Pal = pallidum; pCg = posterior cingulate gyrus; pCO = posterior central operculum; pdPMC = posterior dorsal premotor cortex; pINS = posterior insula; pMFg = posterior middle frontal gyrus; PO = parietal operculum; PP = planum polare; preSMA = pre-supplementary motor area; pSMg = posterior supramarginal gyrus; pSTg = posterior superior temporal gyrus; PT = planum temporale; Put = putamen; SFg = superior frontal gyrus; SMA = supplementary motor area; SPL = superior parietal lobule; splCB = superior posterior lateral cerebellum; spmCB = superior posterior medial cerebellum; Tha = thalamus; TP = temporal pole; vMC = ventral primary motor cortex; VPM = ventral posterior thalamic nucleus; vPMC = ventral premotor cortex; vSC = ventral somatosensory cortex. n.s. = not significant.

effects analysis ($t>4.13$, df = 12, $p_{FDR}<1\%$). The highest normalized effect was located in left ventral somatosensory cortex (MNI *xyz* = [−58, −10, 40]; normalized $\beta=4.29$; see Table 1). Significant increased effects were also observed in pre-supplementary motor area (preSMA) and supplementary motor area (SMA) and bilaterally in ventral motor cortex and auditory cortical areas along the superior temporal gyrus and within the Sylvian fissure. Subcortically, there were increased effects bilaterally in the putamen, pallidum, and thalamus. In the cerebellum, speech-related effects were revealed bilaterally in lobule VI.

The comparison of unexpected somatosensory perturbation in the *absence* of speech with silent observation demonstrated in Fig. 2B revealed significant increased bilateral effects along the precentral gyrus ($t>4.34$, df = 12, $p_{FDR}<1\%$). Compared to the *speech-baseline* contrast, the highest normalized effect size for the *perturbed non-speech–baseline* contrast shifted from left to right hemisphere ventral somatosensory cortex (MNI *xyz* = [62, −8, 36]; normalized $\beta=4.02$). Significant increased effects were also evident within SMA and preSMA, bilateral putamen, and right hemisphere thalamus. In the cerebellum, small clusters of increased effects were found within bilateral superior lateral (lobule VI) and left inferior (lobule VIII) regions.

Fig. 2C demonstrates the normalized group effects for *perturbed speech* relative to *speech* ($t>4.32$, df = 12, $p_{FDR}<1\%$). This contrast characterizes perturbation-related responses in the presence of speech that are not necessarily speech specific. It should be noted that we do not expect or assume that all of the areas involved in somatosensory feedback control would be speech specific (cf., Christoffels et al., 2007). The maximum normalized effect size for the *perturbed speech–speech* contrast was located in right anterior supramarginal gyrus (MNI *xyz* = [62, −20, 26]; normalized $\beta=2.33$). Significant increased effects were noted within preSMA, SMA, and inferior frontal, precentral, and supramarginal gyri. Subcortically, increased effects were observed in thalamus and bilateral inferior posterior cerebellum (lobule VIII).

No voxels demonstrated a response that passed a whole-brain corrected significance threshold of $p_{FDR}<1\%$ for the *perturbed speech–perturbed non-speech* contrast (characterizing speech-related areas in the presence of a perturbation, that are not necessarily *specific* or *sensitive* to processes involved in somatosensory feedback control). However, the ROI analysis did demonstrate significant increased effects for many of the same regions that showed significant increased effects for the *speech–baseline* contrast. Specifically, for the *perturbed speech–perturbed non-speech* contrast significant increased effects were observed for bilateral ventral motor cortex, SMA, preSMA, and auditory cortical regions along the superior temporal gyrus and within the Sylvian fissure (see Supplementary Fig. 4). Subcortically, significant increased effects were found for right hemisphere putamen and anterior medial cerebellum and bilateral pallidum and thalamus.

The interaction in this experimental design [(*perturbed speech–speech*)–(*perturbed non-speech–baseline*)] presents the voxels that show a co-dependence between speech production and perturbation. Two clusters of activity for which the effect of perturbation is greater during speech survived a corrected significance threshold ($t>7.84$, df = 12, $p_{FDR}<1\%$; Supplementary Fig. 2). The clusters were located within right anterior insula (peak MNI *xyz* = [46, 18, −8]; normalized $\beta=1.88$) and left superior frontal gyrus (peak MNI *xyz* = [0, 32, 52]; normalized $\beta=1.70$). While this contrast is informative, indicating that the significant increased effects observed in anterior insula and left superior frontal gyrus are specific to the *perturbed speech* condition, as noted above, we do not assume that all of the areas involved in somatosensory feedback control would be purely speech specific (cf., Christoffels et al., 2007). This contrast may also be limited by the fact that it assumes a large amount of task additivity through the use of multiple subtractions, which has been called under question in the past (e.g., Sidtis et al., 1999).

The standard normalization procedure in voxel-based analyses falls short of ensuring alignment of the structural, and presumably functional, regions across even a small subject cohort (Nieto-Castanon et al., 2003). In particular, two adjacent points across the bank of a major sulcus, separated by less than a millimeter in 3-D volume space, may lie several centimeters apart with respect to their distance along the cortical sheet. Isotropic smoothing in volume space ignores this distinction, blurring responses from two potentially different functional regions. The limitation in spatial sensitivity of voxel-based analysis can be overcome by comparing functional responses within like anatomical regions of interest across subjects. We have developed a parcellation scheme that delimits regions of interest for the cerebral cortex and cerebellum based on an individual subject's neuroanatomical landmarks (See Supplementary Fig. 1; Tourville and Guenther, 2003). This parcellation scheme provides an automatic and standardized method to delineate ROIs that are particularly relevant for neuroimaging studies of speech. We supplemented our voxel-based analyses with region-of interest (ROI) based analyses (Nieto-Castanon et al., 2003). Results from the ROI analysis confirmed our initial hypothesis that the ventral premotor cortex and the inferior frontal gyrus *pars opercularis* were significantly left-lateralized for speech relative to baseline (Fig. 3; $t_{two\text{-}tailed}>2.18$; df = 12; $p<5\%$ for tests on 3 ROIs). Tests of hemispheric differences on the remaining ROIs for the *speech–baseline* contrast yielded a significantly greater left hemisphere effect for the ventral motor cortex and a significantly greater right hemisphere effect for ventral somatosensory cortex, planum temporale, and superior posterior medial cerebellum ($t_{two\text{-}tailed}>3.28$, df = 12, $p_{FDR}<5\%$ for tests on 36 ROIs; Supplementary Fig. 3). Laterality tests for the *perturbed speech–perturbed non-speech* contrast also demonstrated a significantly greater left hemisphere effect for ventral motor cortex ($t_{two\text{-}tailed}>3.40$, df = 12, $p_{FDR}<5\%$ for tests on 36 ROIs; Supplementary Fig. 4).

The normalized effect sizes for cortical regions of the *perturbed non-speech–baseline* and *perturbed speech–speech* contrasts demonstrated a tendency to be right-lateralized (See Supplementary Fig. 5 for the results from the *perturbed non-speech–baseline* contrast). Initial tests of laterality confirmed that the inferior frontal gyrus *pars triangularis*, inferior frontal gyrus *pars opercularis,* and ventral premotor cortex were significantly right-lateralized for the *perturbed speech–speech* contrast ($t_{two\text{-}tailed}>2.18$, df = 12, $p<5\%$ for tests on 3 ROIs; see Fig. 3). Exploratory tests of laterality on the remaining 36 ROIs demonstrated that the planum polare, planum temporale, anterior dorsal and posterior ventral superior temporal sulcus, preSMA, anterior cingulate gyrus, thalamus, and anterior and posterior supramarginal gyrus were significantly right-lateralized ($t_{two\text{-}tailed}>2.79$, df = 12, $p_{FDR}<5\%$ for tests on 36 ROIs; Supplementary Fig. 6). None of the ROIs were significantly left-lateralized in the *perturbed speech–speech* contrast. Finally, we assessed whether the 39 ROIs showed a significant interaction between speech and perturbation. None of the 39 ROIs demonstrated statistically significant increases in effects for this interaction (See Supplementary Fig. 7).

Regional interactions

Structural equation modeling (SEM) was used to empirically evaluate differences in the functional connectivity between the *speech* and *perturbed speech* conditions for a set of five cortical regions of interest hypothesized to be part of the somatosensory feedback control network for speech. According to the DIVA model, somatosensory error signals in bilateral anterior supramarginal gyrus are transformed into compensatory motor commands issued from bilateral ventral motor cortex via connections through right inferior frontal and ventral premotor cortex (Golfinopoulos et al., 2010; Tourville and Guenther, 2010). The increased activity in these areas due to the somatosensory perturbation noted in the *perturbed speech–speech* contrast should therefore be accompanied by increased effective connectivity between bilateral anterior supramarginal

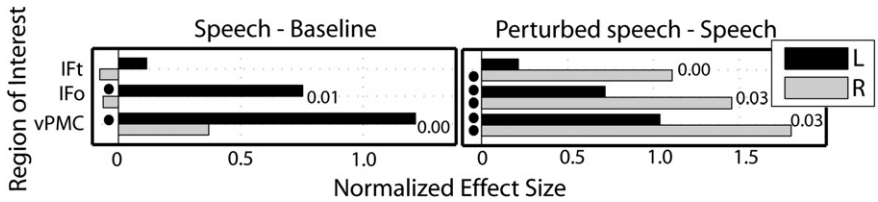

Fig. 3. Normalized effects are shown for those ROIs hypothesized to be left-lateralized for speech under normal conditions, but become right-lateralized when compensatory responses are induced: inferior frontal gyrus *pars triangularis* (IFt), inferior frontal gyrus *pars opercularis* (IFo), ventral premotor cortex (vPMC). Right and left hemisphere responses are indicated by light and dark bars, respectively. Black bullets adjacent to response bars denote significant ROI level effects ($p<5\%$). The p-value is provided for those ROIs that demonstrated significant laterality effects for the *speech–baseline* (left) and the *perturbed speech–speech* (right) contrasts. These initial tests of laterality confirmed that activity within the IFo and vPMC showed a left-over-right asymmetry for the speech network, but along with IFt became significantly right-lateralized for the inferred somatosensory feedback control network ($t_{two\text{-}tailed}>2.18$, df = 12, $p<5\%$ for tests on 3 ROIs).

gyrus and right ventral premotor cortex and right inferior frontal gyrus *pars triangularis* (cf., auditory feedback control network characterized by Tourville et al., 2008). To test this prediction, effective connectivity within the network shown in Fig. 4 for the *perturbed speech* and *speech* conditions was compared. The implied covariance of the model adequately fits the empirical covariance across all conditions ($\chi^2_{uncon}=3.71$, $df_{uncon}=5$, $p_{uncon}=0.59$) and met all goodness-of-fit criteria ($GFI_{uncon}=1.00$, $AGFI_{uncon}=1.00$, $RMR_{uncon}=0.02$, $RMSEA_{uncon}<0.01$).

The unconstrained model, in which connection strengths were allowed to vary across the *perturbed speech* and *speech* conditions, provided a significantly better fit to the data than did the fully constrained null model ($\chi^2_{diff}=26.42$, df = 10, $p<0.01$), indicating differences in global effective connectivity when the perturbation was applied during speech as compared to the *speech* condition. Estimates of the path coefficients along with their corresponding standard errors, z scores, and p-values for both the *perturbed speech* and *speech* conditions are summarized in Table 2. The last two columns of the table list the z scores and associated p-values for pair-wise parameter differences between the two conditions. Pair-wise comparisons of the path coefficients demonstrated that positive connection strengths from left anterior supramarginal gyrus to right anterior supramarginal gyrus and from left anterior supramarginal gyrus to right ventral premotor cortex were significantly greater for *perturbed speech* as compared to *speech*. The reciprocal path connections between the right anterior supramarginal gyrus and the right inferior frontal gyrus *pars triangularis* were also significantly greater for *perturbed speech* relative to *speech*. Likewise, the reciprocal path connections between the right ventral premotor cortex and right ventral primary motor cortex were significantly greater for *perturbed speech* relative to *speech*.

Discussion

In this study we utilized an unexpected perturbation paradigm in order to investigate the neural substrates underlying the somatosensory feedback control of speech. The unpredictability of the perturbation is critical since it prevents subjects from adapting their feedforward commands (or motor programs) over multiple trials to minimize the influence of the perturbation on speech motor output. The dynamic perturbation prompts subjects to rely more directly on the subsystem that monitors somatosensory feedback to determine whether or not it is within the expected range and, when it is not, contribute to the adjustment of speech motor commands to produce the desired speech sound. Voxel-based analysis revealed a peak normalized effect in right anterior supramarginal gyrus for *perturbed speech* relative to *speech*. Increased activity within right supramarginal gyrus has previously been observed in conditions in which there was a discrepancy between expected and afferent somatosensory feedback (Baciu et al., 2000; Jenmalm et al., 2006; Naito et al., 2005; Schmitz et al., 2005). ROI functional analyses, using *a priori* defined anatomical ROIs, demonstrated that effects in anterior supramarginal gyrus were significant in both hemispheres when somatosensory feedback was unexpectedly perturbed during speech relative to speech, but not when speech was produced under unperturbed conditions relative to baseline. Typically, activity in bilateral supramarginal gyrus is not active during normal speech production (for a review see Indefrey and Levelt, 2004). One possible reason for the lack of activity in this region as predicted by the DIVA model is that under normal (unperturbed) conditions somatosensory feedback control contributes little to the speech of fluent adult speakers, since feedforward speech motor programs for frequently produced syllables are highly tuned over the course of development and thus speakers make few, if any, significant somatosensory errors (Golfinopoulos et al., 2010; Guenther, 1994, 1995, 2006; Guenther et al., 2006; Tourville and Guenther, 2010). Recent evidence substantiating this hypothesis comes from an fMRI study performed by Blakemore et al. (1998). The authors demonstrated that activity within bilateral parietal operculum is suppressed during self-generated tactile stimulation as compared to externally produced stimulation. These findings provide support for the theory that an efference copy of motor commands reduces tactile sensation of

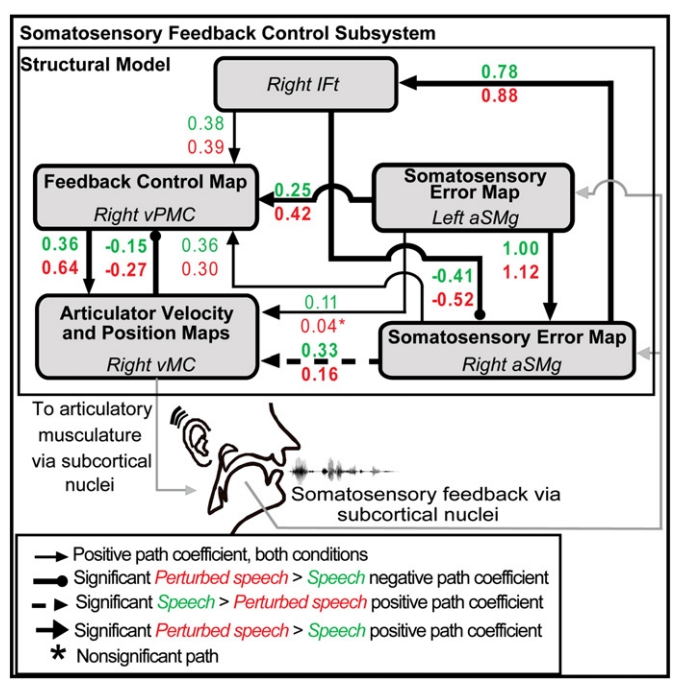

Fig. 4. Schematic of the path diagram evaluated by structural equation modeling for the *perturbed speech* and *speech* conditions. Inter-regional effective connectivity within the network of regions shown was significantly modulated by the jaw perturbation. Path coefficients for all projections shown were significant in both conditions except that from left anterior supramarginal gyrus (aSMg) to right ventral motor cortex (vMC; *perturbed speech* $p=0.60$). Pair-wise comparisons of path coefficients in the two conditions revealed significant interactions due to the somatosensory perturbation in the projections that are highlighted in bold. Abbreviations: IFt = inferior frontal gyrus *pars triangularis*; vPMC = ventral premotor cortex.

Table 2
Effective connectivity determined by structural equation modeling of the network shown in Fig. 4 for the *perturbed speech* and *speech* conditions.

Network path	*Speech*				*Perturbed speech*				*Perturbed speech–Speech*	
	Path coefficient	Standard error	Z	P	Path coefficient	Standard error	Z	P	Z	P
Left aSMg → Right aSMg	1.00	0.02	42.38	<0.001	1.12	0.05	21.25	<0.001	2.18	<0.05
Left aSMg → Right vPMC	0.25	0.02	10.60	<0.001	0.42	0.06	7.08	<0.001	2.75	<0.05
Left aSMg → Right vMC	0.11	0.03	3.60	<0.001	0.04	0.07	0.52	>0.05	−0.94	0.35
Right aSMg → RightvMC	0.33	0.03	11.12	<0.001	0.16	0.06	2.63	0.009	−2.53	<0.05
Right aSMg → Right vPMC	0.36	0.03	12.39	<0.001	0.30	0.06	5.22	<0.001	−1.05	0.29
Right aSMg → Right IFt	0.78	0.02	35.34	<0.001	0.88	0.05	19.55	<0.001	2.09	<0.05
Right IFt → Right aSMg	−0.41	0.03	−13.92	<0.001	−0.52	0.05	−9.79	<0.001	−2.16	<0.05
Right IFt → Right vPMC	0.38	0.02	20.62	<0.001	0.39	0.05	8.26	<0.001	0.22	0.83
Right vPMC → Right vMC	0.36	0.05	7.15	<0.001	0.64	0.08	7.99	<0.001	3.49	<0.05
Right vMC → Right vPMC	−0.15	0.05	−3.24	0.001	−0.27	0.06	−4.29	<0.001	−1.98	<0.05

Significant increases in effective connectivity due to the somatosensory perturbation were found in connection strengths from left aSMg to right aSMg, from left aSMg to right vPMC, from right aSMg to right IFt, from right IFt to right aSMg, from right vPMC to right vMC, and from right vMC to Right vPMC. Abbreviations: aSMg = anterior supramarginal gyrus; IFt = inferior frontal gyrus pars triangularis; vPMC = ventral premotor cortex; vMC = ventral motor cortex.

self-generated stimulation by accurately predicting the sensory consequences of movement (Blakemore et al., 2000). Since this pivotal study, several neuroimaging studies involving speech and non-speech monitoring have demonstrated bilateral engagement of supramarginal gyrus under conditions in which there is a discrepancy between expected and afferent sensory feedback (e.g., Downar et al., 2000; Fink et al., 1999; Hashimoto and Sakai, 2003). The structural equation modeling analysis of the present study demonstrated a significantly greater influence of left anterior supramarginal gyrus on right anterior supramarginal gyrus when somatosensory feedback is unexpectedly perturbed during speech relative to speech. The finding that bilateral anterior supramarginal gyrus are engaged and functionally coupled under conditions in which there is a mismatch between expected and actual somatosensory feedback during speech provides compelling support to the DIVA model's prediction that these regions operate together within a larger neural network to influence speech production when somatosensory feedback is unexpectedly perturbed during speech.

In addition to increased right-lateralized supramarginal gyrus activity, we observed increased effects in bilateral inferior posterior cerebellum (lobule VIII) when somatosensory feedback was unexpectedly perturbed during speech relative to speech. Increased activation in this inferior posterior region of the cerebellum along with more superior fields has been associated with non-speech tongue and lip movements (Grodd et al., 2001). Cerebellar lobule VIII has been associated with the timing of bimanual complex movements (Habas et al., 2004; Habas and Cabanis, 2006) and discrete index finger movements, where somatosensory feedback can influence movement transitions (Habas and Cabanis, 2008) as well as for compensatory responses to unexpected execution errors (Diedrichsen et al., 2005). In speech production, increased lobule VIII activation has been associated with the *overt* production of repeated syllables (Riecker et al., 2005, 2006), syllable sequences (Bohland and Guenther, 2006), monosyllables (Ghosh et al., 2008), and compensatory responses to shifts in auditory feedback (Tourville et al., 2008), as well as *covert* singing production relative to *covert* speech production (Callan et al., 2006). Although the functional recruitment of cerebellar lobule VIII in speech motor control remains largely under speculation, we suggest that this region may facilitate sensory cortical areas in the monitoring and/or adjustment of articulator movements when sensory feedback is unexpectedly perturbed during speech.

Recent fMRI studies in our laboratory have demonstrated a leftward asymmetry in both the posterior inferior frontal gyrus and ventral premotor cortex for speech under normal (unperturbed) conditions (Ghosh et al., 2008; Tourville et al., 2008). This trend is supported by the *speech–baseline* contrast of the present study, where effects were left-lateralized in the inferior frontal gyrus *pars opercularis* and ventral premotor cortex. When compared to speech, the unexpected balloon inflation during speech that prompts subjects to adjust their planned articulator trajectories resulted in right-lateralized activity within precentral and inferior frontal gyrus. Baciu et al. (2000) also reported increased activation within right inferior frontal gyrus *pars triangularis* in association with compensatory responses to static perturbations of the lips during the articulation of a vowel that requires lip protrusion. In addition, right inferior frontal gyrus engagement has previously been implicated for conditions in which monitoring demands are increased due to a conflict between sensory expectations and feedback (Downar et al., 2000; Fink et al., 1999; Naito et al., 2005; Schmitz et al., 2005) and when visually observing orofacial movements (Buccino et al., 2001). Collectively, these studies suggest that right inferior frontal gyrus maintains sensorimotor representations and may be involved in the monitoring and/or adjustment of ongoing speech movements. Support for this hypothesis was demonstrated by the structural equation modeling analysis in the present study. Compared to *speech*, the *perturbed speech* condition involved a significantly greater *reciprocal* influence between right anterior supramarginal gyrus and right inferior frontal gyrus *pars triangularis*. These results suggest that projections between right higher-order somatosensory cortical and inferior frontal regions are critical for distinguishing between *perturbed speech* and *speech* conditions and may relay sensory error induced by unexpected somatosensory perturbation to right inferior frontal gyrus to facilitate sensory feedback monitoring and/or the correction of speech movements.

The right-lateralized contribution from ventral premotor cortex associated with *perturbed speech* relative to *speech* differentiates our study from the earlier fMRI experiment on somatosensory feedback perturbation during speech in which no lateral premotor cortex activity was reported in association with compensatory responses during speech (Baciu et al., 2000). However, major differences between the study designs including the use of a static somatosensory feedback perturbation in which subjects may have adapted their feedforward motor commands to compensate make it very difficult to interpret the lack of activation in this region in the earlier study. The results of several neuroimaging studies do support our findings of right-lateralized contributions from ventral premotor cortex for the generation of motor responses to unexpected changes in sensory feedback (Grafton et al., 2008; Houde and Nagarajan, 2007; Tourville et al., 2008). Findings from the non-human primate literature indicate that ventral premotor cortex receives projections from parietal area 7b (Cavada and Goldman-Rakic, 1989; Neal et al., 1990), a region thought to be the homologous region of human supramarginal gyrus (Aboitiz and García, 1997), and projects to primary motor cortex (Barbas and Pandya, 1987). Projections such as these provide a substrate for sensory error during speech movements to influence speech motor plans (cf., Watkins et al., 2008). In the present study,

structural equation modeling demonstrated a significant increased influence between left anterior supramarginal gyrus and right ventral premotor cortex as well as a significantly greater reciprocal influence between right ventral premotor cortex and right ventral motor cortex for *perturbed speech* relative to *speech*. Taken together, these results suggest that right-lateralized ventral premotor cortex is well situated to be at the heart of a sensory motor circuit responsible for translating sensory errors into corrective motor commands when somatosensory feedback is unexpectedly perturbed during speech production.

Increased activity of right-lateralized prefrontal and premotor cortices was accompanied by increased bilateral ventral motor cortical activity in the *perturbed speech–speech* contrast. Activity in bilateral ventral motor cortex is critical for speech production (for a review see Indefrey and Levelt, 2004), but is often shown to be left-lateralized during overt speech production (e.g., Ghosh et al., 2008; Tourville et al., 2008). Additional support for this trend comes from the present study in which we found a significantly greater left hemisphere effect for ventral motor cortex in the *speech–baseline* contrast. The ROI functional analyses demonstrated that this left hemisphere asymmetry in ventral motor cortex observed for the *speech–baseline* contrast does not persist for the *perturbed speech–speech* contrast. Bilateral (without significant laterality) recruitment of ventral motor cortex has previously been demonstrated for compensatory responses to auditory-feedback-based perturbations during speech (Tourville et al., 2008). Direct evidence for left primary motor cortical involvement during dynamic somatosensory-feedback based perturbations of speech comes from a recent transcranial magnetic stimulation study by Ito et al. (2005) implicating the region in upper lip compensatory responses to unexpected jaw perturbations.

The *perturbed speech–speech* contrast also demonstrated increased effects in bilateral anterior insula, and right hemisphere anterior cingulate gyrus and pre-supplementary motor area (preSMA). Increased anterior cingulate cortex and bilateral insula recruitment is similar to the pattern of activity observed in a recent fMRI study on verbal self-monitoring (Christoffels et al., 2007). Christoffels et al. (2007) propose that speech monitoring may largely rely on non-language specific areas such as the insular and anterior cingulate cortices that have previously been implicated in general performance monitoring (e.g., Botvinick et al., 2001; Fiehler et al., 2004). In the present study, the peak effect in anterior cingulate cortex (MNI $xyz = 4\ -3\ 40$) was located in the right hemisphere posterior to the vertical plane passing through the anterior commissure (VAC). A second peak (MNI $xyz = 2\ 16\ 22$) was found anterior to the VAC. While activity in the more posterior region is generally observed during simple motor tasks and somatosensory stimulation (Picard and Strick, 1996), activity in the anterior region has been implicated in more complex aspects of motor behavior including the monitoring of response errors (for a review see Ridderinkhof et al., 2004). Christoffels et al. (2007) propose that projections from bilateral insula to the anterior cingulate facilitate speech monitoring by conveying contextual information about the sensory state of the periphery. This theory suggests that the significant increased effects in bilateral anterior insula and right anterior cingulate gyrus for the *perturbed speech–speech* contrast reflect increased monitoring of speech output. An alternative (though not necessarily incompatible) proposal for right inferior frontal (including inferior frontal gyrus *pars opercularis* and insula) and preSMA activation comes from a recent fMRI study by Xue et al. (2008) that implicated these regions in a general (non-language specific) mechanism for inhibitory control. Xue et al. (2008) demonstrated that successful *inhibition* of speech and manual motor acts involves increased activation of right inferior frontal regions and preSMA. In light of these recent findings, the increased effects in right inferior frontal and preSMA regions in the *perturbed speech–speech* contrast may reflect a common strategy among subjects to inhibit planned jaw movements when somatosensory feedback is unexpectedly perturbed. It should be noted that the right anterior insula was also implicated in the interaction contrast (*perturbed speech–speech*)–(*perturbed non-speech–baseline*), indicating that the effects observed in anterior insula show a co-dependence between speech production and perturbation, which is more consistent with the proposal of an increased verbal monitoring role for anterior insula during the *perturbed speech* condition (cf., Christoffels et al., 2007). However, since the present study did not investigate the neural responses associated with somatosensory error monitoring independently of error correction during speech, it remains difficult to explicitly assign functional roles to the pre-supplementary motor area, insular and anterior cingulate corticies on the basis of this study alone.

For more than a century researchers have implicated a role for left posterior inferior frontal and/or ventral premotor regions in the production of speech (Broca, 1861; Penfield and Rasmussen, 1950; Penfield and Roberts, 1959; Raichle et al., 1994; Riecker et al., 2000; Sidtis et al., 2003; Xue et al., 2008; Tourville et al., 2008; Ghosh et al., 2008; Papoutsi et al., 2009). Damage to these regions is often associated with apraxia of speech (e.g., Hillis et al., 2004; Robin et al., 2007), a speech motor disorder characterized by an inability to properly access speech motor programs in the absence of morphological defect of the speech musculature. A notable finding from the present study was a general greater dependency on right hemisphere inferior frontal and premotor regions when compensation was prompted by the unexpected inflation of the balloon during speech relative to speech. Increased right hemisphere engagement of inferior frontal and/or lateral premotor cortical regions has been associated with conditions in which compensatory responses were prompted by sensory feedback error in other neuroimaging studies of speech (e.g., Baciu et al., 2000; Houde and Nagarajan, 2007; Toyomura et al., 2007; Tourville et al., 2008). We propose that left-lateralized inferior frontal and lateral premotor cortex contributes to speech movements by using well-learned (feedforward) speech motor programs, while right-lateralized inferior frontal and lateral premotor cortex contributes to speech movements by using sensory feedback. Such a theory is consistent with reports from clinical studies that left hemisphere damaged aphasic and apraxia of speech subjects are largely spared in terms of their ability to compensate for somatosensory perturbations during speech (Baum et al., 1997; Baum, 1999; Jacks, 2008). This theory also explains why damage to right hemisphere premotor areas typically does not result in apraxia of speech (Duffy, 2005), since, according to this theory, feedforward speech motor programs are typically maintained in the left hemisphere.

Several methodological issues regarding the current study should be acknowledged. First, direct verification of behavioral performance during the scanning session was not possible due to current limitations in technology that prevent the measuring of the kinematics of the articulators inside the MR scanner. In addition, due to a technical malfunction, audio recordings of subjects' productions inside the scanner were lost, preventing acoustic analysis that could confirm whether subjects' acoustic responses to the somatosensory perturbation met the acoustic requirements of speech sound targets. It should be noted, however, that interpretation of the data from the present study are not dependent on the subject achieving complete compensation in response to the perturbation; instead it requires only that the somatosensory feedback control network for speech be invoked by the unexpected somatosensory feedback perturbation (i.e. that the somatosensory subsystem detected that the jaw was no longer moving in the normal way and attempted to adjust the kinematics of functionally relevant but unimpeded articulators to respond to the sudden perturbation). Our pilot articulometry results as well as the results of a large number of previous studies have shown the ability of subjects to compensate for jaw perturbations (e.g., Folkins and Abbs, 1975; Gay et al., 1981; Gomi et al., 2002; Ito et al., 2005; Kelso et al., 1984; Lane et al., 2005; Lindblom et al., 1979; Nasir and Ostry, 2006, 2008, 2009; Shaiman,

1989; Tremblay et al., 2003). Moreover, several of these earlier studies have implicated subjects' use of somatosensory feedback to generate compensatory responses to jaw perturbations (e.g., Lindblom et al., 1979; Tremblay et al., 2003; Nasir and Ostry, 2006, 2008, 2009), lending support to the claim that the somatosensory feedback control network was invoked by the unexpected jaw perturbation in this study. Second, like many earlier studies examining the neural correlates of speech monitoring (e.g., Baciu et al., 2000; Fu et al., 2006; Hashimoto and Sakai, 2003; McGuire et al., 1996; Tourville et al., 2008; Toyomura et al., 2007) an external stimulus was introduced in the process of investigating the neural bases of somatosensory feedback control. As Christoffels et al. (2007) have noted, when an external stimulus is introduced it is not entirely clear whether the associated neural findings reflect normal speech monitoring. This is an intriguing technical challenge that, as yet, has not been completely solved in terms of the somatosensory feedback control of speech production. It is worth noting, however, that the highest effect for the *perturbed speech–speech* contrast was considerably more posterior than that for the *perturbed non-speech–baseline* contrast, suggesting that additional mechanisms beyond those directly associated with the somatosensory stimulus are recruited when the stimulus is applied during speech, consistent with our interpretation. Finally, with the current study design, we cannot fully eliminate the possibility that auditory feedback also influenced motor responses during the perturbed feedback condition. Although auditory feedback was largely suppressed since subjects wore ear plugs throughout the experiment, it is possible that subjects still perceived some acoustic effects of the perturbation through residual hearing ability or bone conduction. Indeed, the ROI analysis demonstrated significant effects in right-lateralized planum temporale and in right posterior superior temporal gyrus when somatosensory feedback was unexpectedly perturbed during speech relative to speech production. Increased activity within right posterior superior temporal gyrus was also reported in the earlier somatosensory perturbation experiment in association with compensating for the static lip tube (Baciu et al., 2000). In this earlier study subjects were only instructed to articulate, without voicing, the vowel stimulus, which suggests that activity in right posterior superior temporal gyrus was not directly associated with speakers hearing their own (incorrect) auditory feedback. Further study is necessary to determine whether higher-order auditory cortical activation when somatosensory feedback is unexpectedly perturbed during speech reflects concomitant auditory feedback control and/or the influence of somatosensory stimulation on auditory association areas (for a review see Zheng, 2009).

Conclusion

The present investigation highlights the recruitment of cerebellar, motor, and fronto-parietal regions and a generally greater contribution from right hemisphere cortical regions for *perturbed speech* relative to *speech*. These findings augment the currently scant imaging data on the neural substrates underlying somatosensory feedback control during speech. Voxel-based and ROI-based analyses indicated that unexpected perturbation of somatosensory feedback during speech compared to speech is associated with bilateral responses in anterior supramarginal gyrus, with a somewhat larger response in the right hemisphere. This supports the DIVA model prediction that under conditions in which somatosensory feedback is not within the expected range during speech movements, cells within bilateral supramarginal gyrus (the location of the model's *somatosensory error map*) are highly active (Golfinopoulos et al., 2010; Guenther, 2006; Guenther et al., 2006; Tourville and Guenther, 2010). The contrast of *perturbed speech* with *speech* also revealed increased activity within right-lateralized ventral premotor cortex and inferior frontal gyrus. Structural equation modeling revealed a significant increased influence from left anterior supramarginal gyrus to right anterior supramarginal gyrus and from left anterior supramarginal gyrus to right ventral premotor cortex as well as a significant increased reciprocal influence between right ventral premotor cortex and right ventral motor cortex and right anterior supramarginal gyrus and right inferior frontal gyrus *pars triangularis* for *perturbed speech* relative to *speech*. These results suggest that bilateral anterior supramarginal gyrus, right inferior frontal gyrus *pars triangularis*, right ventral premotor and motor cortices are functionally coupled and contribute to adjustments in speech motor output when somatosensory feedback is unexpectedly perturbed during speech.

Supplementary materials related to this article can be found online at doi:10.1016/j.neuroimage.2010.12.065.

Acknowledgments

This research was supported by the National Institute on Deafness and other Communication Disorders (R01 DC02852, F. Guenther PI) and by CELEST, an NSF Science of Learning Center (NSF SMA-0835976). Imaging was made possible with support from the National Center for Research Resources grant P41RR14075 and the MIND Institute. The authors would like to thank Joseph Perkell, Mark Tiede, Kevin Reilly, Oren Civier, Bruce Fischl, Julie Goodman, Lawrence Wald and the anonymous reviewers for their time and consideration.

References

Abbs, J.H., Gracco, V.L., 1984. Control of complex motor gestures: orofacial muscle responses to load perturbation of lip during speech. J. Neurophysiol. 51, 705–723.
Aboitiz, F., García, V., 1997. The evolutionary origin of the language areas in the human brain. A neuroanatomical perspective. Brain Res. Rev. 25, 381–396.
Au Duong, M.V., Boulanouar, K., Audoin, B., Treseras, S., Ibarrola, D., Malikova, I., Confort-Gouny, S., Celsis, P., Pelletier, J., Cozzone, P.J., Ranjeva, J.P., 2005. Modulation of effective connectivity inside the working memory network in patients at the earliest stage of multiple sclerosis. Neuroimage 24, 533–538.
Baciu, M., Abry, C., Segebarth, C., 2000. Equivalence motrice et dominance hémisphérique. Le cas de la voyelle (u). Étude IRMf. Actes JEP 23, 213–216.
Barbas, H., Pandya, D.N., 1987. Architecture and frontal cortical connections of the premotor cortex (area 6) in the rhesus monkey. J. Comp. Neurol. 256, 211–228.
Baum, S.R., 1999. Compensation for jaw fixation by aphasic patients under conditions of increased articulatory demands: a follow-up study. Aphasiology 13, 513–527.
Baum, S.R., McFarland, D.H., Diab, M., 1996. Compensation to articulatory perturbation: perceptual data. J. Acoust. Soc. Am. 99, 3791–3794.
Baum, S.R., Kim, J.A., Katz, W.F., 1997. Compensation for Jaw Fixation by Aphasic Patients. Brain Lang 56, 354–376.
Belin, P., Zatorre, R., Hoge, R., Evans, A., Pike, B., 1999. Event-related fMRI of the auditory cortex. Neuroimage 10, 417–429.
Birn, R.M., Bandettini, P.A., Cox, R.W., Shaker, R., 1999. Event-related fMRI of tasks involving brief motion. Hum. Brain Mapp. 7, 106–114.
Blakemore, S.J., Wolpert, D.M., Frith, C.D., 1998. Central cancellation of self-produced tickle sensation. Nat. Neurosci. 1, 635–640.
Blakemore, S.J., Wolpert, D., Frith, C., 2000. Why can't you tickle yourself? NeuroReport 11 (11), R11–R16.
Bohland, J.W., Guenther, F.H., 2006. An fMRI investigation of syllable sequence production. Neuroimage 32, 821–841.
Botvinick, M., Braver, T., Barch, D., Carter, C., Cohen, J., 2001. Conflict monitoring and cognitive control. Psychol. Rev. 108, 624–652.
Broca, P., 1861. Remarques sur la siège de la faculté de la parole articulée, suivies d'une observation d'aphémie perte de parole. Bull. Soc. Anat. 36, 330–357.
Buccino, G., Binkofski, F., Fink, G.R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R.J., Zilles, K., Rizzolatti, G., Freund, H.-J., 2001. Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. Eur. J. Neurosci. 13, 400–404.
Büchel, C., Friston, K.J., 1997. Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modeling and fMRI. Cereb. Cortex 7, 768–778.
Callan, D.E., Tsytsarev, V., Hanakawa, T., Callan, A.M., Katsuhara, M., Fukuyama, H., Turner, R., 2006. Song and speech: brain regions involved with perception and covert production. Neuroimage 31, 1327–1342.
Cavada, C., Goldman-Rakic, P.S., 1989. Posterior parietal cortex in rhesus monkey: II. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe. J. Comp. Neurol. 287, 422–445.
Christoffels, I.K., Formisano, E., Schiller, N.O., 2007. Neural correlates of verbal feedback processing: an fMRI study employing overt speech. Hum. Brain Mapp. 28, 868–879.
Chumbley, J.R., Friston, K.J., 2009. False discovery rate revisited: FDR and topological inference using Gaussian random fields. Neuroimage 44, 62–70.
Dale, A., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage 9, 179–194.

Diedrichsen, J., Hashambhoy, Y., Rane, T., Shadmehr, R., 2005. Neural correlates of reach errors. J. Neurosci. 25, 9919–9931.
Downar, J., Crawley, A.P., Mikulis, D.J., Davis, K.D., 2000. A multimodal cortical network for the detection of changes in the sensory environment. Nat. Neurosci. 3, 277–283.
Duffy, J.R., 2005. Motor Speech Disorders: Substrates, Differential Diagnosis, and Management, 2nd ed. Elsevier, St. Louis (MO).
Evans, A.C., Collins, D.L., Mills, S.R., Brown, E.D., Kelly, R.L., Peters, T.M., 1993. 3D statistical neuroanatomical models from 305 MRI volumes. Proc. IEEE Nucl. Sci. Symp. Med. Imaging 3, 1813–1817.
Fiehler, K., Ullsperger, M., von Cramon, D.Y., 2004. Neural correlates of error detection and error correction, is there a common neuroanatomical substrate? Eur. J. Neurosci. 9, 3081–3087.
Fink, G.R., Marshall, J.C., Halligan, P.W., Frith, C.D., Driver, J., Frackowiak, R.S., Dolan, R.J., 1999. The neural consequences of conflict between intention and the senses. Brain 122, 497–512.
Fischl, B., Sereno, M.I., Dale, A.M., 1999. Cortical surface-based analysis II. Inflation, flattening, and a surface-based coordinate system. Neuroimage 9, 195–207.
Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M., 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 33, 341–355.
Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale, A.M., 2004. Automatically parcellating the human cerebral cortex. Cereb. Cortex 14, 11–22.
Folkins, J.W., Abbs, J.H., 1975. Lip and jaw motor control during speech: responses to resistive loading of the jaw. J. Speech Lang. Hear. Res. 18, 207–220.
Folkins, J.W., Zimmerman, G.N., 1982. Lip and jaw interaction during speech: responses to perturbation of lower-lip movement prior to bilabial closure. J. Acoust. Soc. Am. 71, 1225–1233.
Fu, C.H., Vythelingum, G.N., Brammer, M.J., Williams, S.C., Amaro Jr., E., Andrew, C.M., Yágüez, L., van Haren, N.E., Matsumoto, K., McGuire, P.K., 2006. An fMRI study of verbal self-monitoring: neural correlates of auditory verbal feedback. Cereb. Cortex 16, 969–977.
Gay, T., Lindblom, B., Lubker, J., 1981. Production of bite-block vowels: acoustic equivalence by selective compensation. J. Acoust. Soc. Am. 69, 802–810.
Genovese, C., Lazar, N., Nichols, T., 2002. Thresholding of statistical maps in neuroimaging using the false discovery rate. Neuroimage 15, 870–878.
Ghosh, S.S., Tourville, J.A., Guenther, F.H., 2008. A neuroimaging study of premotor lateralization and cerebellar involvement in the production of phonemes and syllables. J. Speech Lang. Hear. Res. 51, 1183–1202.
Golfinopoulos, E., Tourville, J.A., Guenther, F.H., 2010. The integration of large-scale neural network modeling and functional brain imaging in speech motor control. Neuroimage 52, 862–874.
Gomi, H., Ito, T., Murano, E.Z., Honda, M., 2002. Compensatory articulation during bilabial fricative production by regulating muscle stiffness. J. Phon. 30, 261–279.
Gonçalves, M.S., Hall, D.A., Johnsrude, I.S., Haggard, M.P., 2001. Can meaningful effective connectivities be obtained between auditory cortical regions? Neuroimage 14, 1353–1360.
Gracco, V.L., Abbs, J.H., 1985. Dynamic control of the perioral system during speech: kinematic analyses of autogenic and nonautogenic sensorimotor processes. J. Neurophysiol. 54, 418–432.
Grafton, S.T., Sutton, J., Couldwell, W., Lew, M., Waters, C., 1994. Network analysis of motor system connectivity in Parkinson's disease: modulation of thalamocortical interactions after pallidotomy. Hum. Brain Mapp. 2, 45–55.
Grafton, S.T., Schmitt, P., Horn, J.V., Diedrichsen, J., 2008. Neural substrates of visuomotor learning based on improved feedback control and prediction. Neuroimage 39, 1383–1395.
Grodd, W., Hulsmann, E., Lotze, M., Wildgruber, D., Erb, M., 2001. Sensorimotor mapping of the human cerebellum: fMRI evidence of somatotopic organization. Hum. Brain Mapp. 13, 55–73.
Guenther, F.H., 1994. A neural network model of speech acquisition and motor equivalent speech production. Biol. Cybern. 72, 43–53.
Guenther, F.H., 1995. Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. Psychol. Rev. 102, 594–621.
Guenther, F.H., 2006. Cortical interactions underlying the production of speech sounds. J. Commun. Disord. 5, 350–365.
Guenther, F.H., Vladusich, T., in press. A neural theory of speech acquisition and production. J. Neurolinguistics.
Guenther, F.H., Ghosh, S.S., Tourville, J.A., 2006. Neural modeling and imaging of the cortical interactions underlying syllable production. Brain Lang. 96, 280–301.
Habas, C., Cabanis, E.A., 2006. Cortical areas functionally linked with the cerebellar second homunculus during out-of-phase bimanual movements. Neuroradiology 48, 273–279.
Habas, C., Cabanis, E.A., 2008. Neural correlates of simple unimanual discrete and continuous movements: a functional imaging study at 3 T. Neuroradiology 50, 367–375.
Habas, C., Axelrad, H., Nguyen, T.H., Cabanis, E.A., 2004. Specific neocerebellar activation during out-of-phase bimanual movements. NeuroReport 15, 595–599.
Hall, D.A., Haggard, M.P., Akeroyd, M.A., Palmer, A.R., Summerfield, A.Q., Elliot, M.R., Gurney, E.M., Bowtell, R.W., 1999. 'Sparse' temporal sampling in auditory fMRI. Hum. Brain Mapp. 7, 213–223.
Hashimoto, Y., Sakai, K.L., 2003. Brain activations during conscious self-monitoring of speech production with delayed auditory feedback: an fMRI study. Hum. Brain Mapp. 20, 22–28.
Hillis, A.E., Work, M., Barker, P.B., Jacobs, M.A., Breese, E.L., Maurer, K., 2004. Re-examining the brain regions crucial for orchestrating speech articulation. Brain 127 (Pt. 7), 1479–1487.
Horwitz, B., McIntosh, A.R., Haxby, J.V., Furey, M., Salerno, J.A., Schapiro, M.B., Rapoport, S.I., Grady, C.L., 1995. Network analysis of PET-mapped visual pathways in Alzheimer type dementia. NeuroReport 6, 2287–2292.
Houde, J.F., Nagarajan, S.S., 2007. How is auditory feedback processed during speaking? Program of the 154th Meeting of the Acoustical Society of America. J. Acoust. Soc. Am. 122 (5, Pt. 2), 3087 New Orleans, Louisiana.
Indefrey, P., Levelt, W.J., 2004. The spatial and temporal signatures of word production components. Cognition 92, 101–144.
Ito, T., Kimura, T., Gomi, H., 2005. The motor cortex is involved in reflexive compensatory adjustment of speech articulation. NeuroReport 16, 1791–1794.
Jacks, A., 2008. Bite block vowel production in apraxia of speech. J. Speech Lang. Hear. Res. 51, 898–913.
Jenmalm, P., Schmitz, C., Forssberg, H., Ehrsson, H.H., 2006. Lighter or heavier than predicted: neural correlates of corrective mechanisms during erroneously programmed lifts. J. Neurosci. 26, 9015–9021.
Kelso, J.A.S., Tuller, B., Vatikiotis-Bateson, E., Fowler, C.A., 1984. Functionally specific articulatory cooperation following jaw perturbations during speech: evidence for coordinative structures. J. Exp. Psychol. Hum. Percept. Perform. 10, 812–832.
Lancaster, J.L., Tordesillas-Gutierrez, D., Martinez, M., Salinas, F., Evans, A., Zilles, K., Mazziotta, J.C., Fox, P.T., 2007. Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. Hum. Brain Mapp. 28, 1194–1205.
Lane, H., Denny, M., Guenther, F.H., Matthies, M.L., Menard, L., Perkell, J.S., Stockmann, E., Tiede, M., Vick, J., Zandipour, M., 2005. Effects of bite blocks and hearing status on vowel production. J. Acoust. Soc. Am. 118, 1636–1646.
Lindblom, B., Lubker, J., Gay, T., 1979. Formant frequencies of some fixed-mandible vowels and a model of speech motor programming by predictive simulation. J. Phon. 7, 147–161.
McFarland, D.H., Baum, S.R., 1995. Incomplete compensation to articulatory perturbation. J. Acoust. Soc. Am. 97, 1865–1873.
McGuire, P.K., Silbersweig, D.A., Frith, C.D., 1996. Functional neuroanatomy of verbal self-monitoring. Brain 119, 907–917.
Naito, E., Roland, P.E., Grefkes, C., Choi, H.J., Eickhoff, S., Geyer, S., Zilles, K., Ehrsson, H.H., 2005. Dominance of the right hemisphere and role of area 2 in human kinesthesia. J. Neurophysiol. 93, 1020–1034.
Nasir, S.M., Ostry, D.J., 2006. Somatosensory precision in speech production. Curr. Biol. 16, 1918–1923.
Nasir, S.M., Ostry, D.J., 2008. Speech motor learning in profoundly deaf adults. Nat. Neurosci. 11, 1217–1222.
Nasir, S.M., Ostry, D.J., 2009. Auditory plasticity and speech motor learning. Proc. Natl. Acad. Sci. USA 106, 20470–20475.
Neal, J.W., Pearson, R.C., Powell, T.P., 1990. The ipsilateral cortico-cortical connections of area 7 with the frontal lobe in the monkey. Brain Res. 509, 31–40.
Nieto-Castanon, A., Ghosh, S.S., Tourville, J.A., Guenther, F.H., 2003. Region of interest based analysis of functional imaging data. Neuroimage 19, 1303–1316.
Papoutsi, M., de Zwart, J.A., Jansma, J.M., Pickering, M.J., Bednar, J.A., Horwitz, B., 2009. From phonemes to articulatory codes: an fMRI study of the role of Broca's area in speech production. Cereb. Cortex 19 (9), 2156–2165.
Penfield, W., Rasmussen, T., 1950. The Cerebral Cortex of Man. A Clinical Study of Localization of Function. Macmillan, New York.
Penfield, W., Roberts, L., 1959. Speech and Brain Mechanisms. Princeton University Press, Princeton.
Penny, W.D., Stephan, K.E., Mechelli, A., Friston, K.J., 2004. Modelling functional integration: a comparison of structural equation and dynamic causal models. Neuroimage 23, S264–S274.
Picard, N., Strick, P.L., 1996. Motor areas of the medial wall: a review of their location and functional activation. Cereb. Cortex 6 (3), 342–353.
Raichle, M., Fiez, J., Videen, T., MacLeod, A., Pardo, J., Fox, P., Petersen, S., 1994. Practice-related changes in human brain functional anatomy during nonmotor learning. Cereb. Cortex 4, 8–26.
Ridderinkhof, K.R., Ullsperger, M., Crone, E.A., Nieuwenhuis, S., 2004. The role of the medial frontal cortex in cognitive control. Science 306, 443–447.
Riecker, A., Ackermann, H., Wildgruber, D., Dogil, G., Grodd, W., 2000. Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. NeuroReport 11, 1997–2000.
Riecker, A., Mathiak, K., Wildgruber, D., Erb, M., Hertrich, I., Grodd, W., Ackermann, H., 2005. fMRI reveals two distinct cerebral networks subserving speech motor control. Neurology 64, 700–706.
Riecker, A., Kassubek, J., Gröschel, K., Grodd, W., Ackermann, H., 2006. The cerebral control of speech tempo: opposite relationship between speaking rate and BOLD signal changes at striatal and cerebellar structures. Neuroimage 29, 46–53.
Robin, D.A., Jacks, A., Ramage, A.E., 2007. The neural substrates of apraxia of speech as uncovered by brain imaging: a critical review. In: Ingham, R.J. (Ed.), Neuroimaging in Communication Sciences and Disorders. Plural Publishing Inc., San Diego, CA, pp. 129–154.
Rowe, J., Stephan, K.E., Friston, K., Frackowiak, R., Lees, A., Passingham, R., 2002. Attention to action in Parkinson's disease: impaired effective connectivity among frontal cortical regions. Brain 125, 276–289.
Schlösser, R.G., Gesierich, T., Kaufmann, B., Vucurevic, G., Hunsche, S., Gawehn, J., Stoeter, P., 2003. Altered effective connectivity during working memory performance in schizophrenia: a study with fMRI and structural equation modeling. Neuroimage 19, 751–763.
Schlösser, R.G., Wagner, G., Sauer, H., 2006. Assessing the working memory network: studies with functional magnetic resonance imaging and structural equation modeling. Neuroscience 139, 91–103.

Schmitz, C., Jenmalm, P., Ehrsson, H.H., Forssberg, H., 2005. Brain activity during predictable and unpredictable weight changes when lifting objects. J. Neurophysiol. 93, 1498–1509.
Shaiman, S., 1989. Kinematic and electromyographic responses to perturbation of the jaw. J. Acoust. Soc. Am. 86, 78–88.
Sidtis, J.J., Strother, S.C., Anderson, J.R., Rottenberg, D.A., 1999. Are brain functions really additive? Neuroimage 5, 490–496.
Sidtis, J.J., Strother, S.C., Rottenberg, D.A., 2003. Predicting performance from functional imaging data: methods matter. Neuroimage 20, 615–624.
Tourville, J.A., Guenther, F.H., 2003. A cortical and cerebellar parcellation system for speech studies. Tech Rep CAS/CNS-03-022. Boston University, Boston, MA.
Tourville, J.A., Guenther, F.H., 2010. The DIVA model: a neural theory of speech acquisition and production. Lang. Cogn. Proc.
Tourville, J.A., Reilly, K.J., Guenther, F.H., 2008. Neural mechanisms underlying auditory feedback control of speech. Neuroimage 39, 1429–1443.
Toyomura, A., Koyama, S., Miyamaoto, T., Terao, A., Omori, T., Murohashi, H., Kuriki, S., 2007. Neural correlates of auditory feedback control in human. Neuroscience 146, 499–503.
Tremblay, S., Shiller, D.M., Ostry, D.J., 2003. Somatosensory basis of speech production. Nature 423, 866–869.
Watkins, K.E., Smith, S.M., Davis, S., Howell, P., 2008. Structural and functional abnormalities of the motor system in developmental stuttering. Brain 131, 50–59.
Xue, G., Aron, A.R., Poldrack, R.A., 2008. Common neural substrates for inhibition of spoken and manual responses. Cereb. Cortex 18, 1923–1932.
Yang, Y., Engelien, A., Engelien, W., Xu, S., Stern, E., Silbersweig, D.A., 2000. A silent event-related functional MRI technique for brain activation studies without interference of scanner acoustic noise. Magn. Reson. Med. 43, 185–190.
Zheng, Z.Z., 2009. The functional specialization of the planum temporale. J. Neurophysiol. 102, 3079–3081.