



Published in final edited form as:

J Speech Lang Hear Res. 2008 October ; 51(5): 1183–1202. doi:10.1044/1092-4388(2008/07-0119).

A Neuroimaging Study of Premotor Lateralization and Cerebellar Involvement in the Production of Phonemes and Syllables

Satrajit S. Ghosh,

Massachusetts Institute of Technology, Cambridge, MA, and Boston University

Jason A. Tourville, and

Boston University

Frank H. Guenther

Boston University, Massachusetts Institute of Technology, and Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA

Abstract

Purpose—This study investigated the network of brain regions involved in overt production of vowels, monosyllables, and bisyllables to test hypotheses derived from the Directions Into Velocities of Articulators (DIVA) model of speech production (Guenther, Ghosh, & Tourville, 2006). The DIVA model predicts left lateralized activity in inferior frontal cortex when producing a single syllable or phoneme and increased cerebellar activity for consonant–vowel syllables compared with steady-state vowels.

Method—Sparse sampling functional magnetic resonance imaging (fMRI) was used to collect data from 10 right-handed speakers of American English while producing isolated monosyllables (e.g., “ba,” “oo”). Data were analyzed using both voxel-based and participant-specific anatomical region-of-interest–based techniques.

Results—Overt production of single monosyllables activated a network of brain regions, including left ventral premotor cortex, left posterior inferior frontal gyrus, bilateral supplementary motor area, sensorimotor cortex, auditory cortex, thalamus, and cerebellum. Paravermal cerebellum showed greater activity for consonant-vowel syllables compared to vowels.

Conclusions—The finding of left-lateralized premotor cortex activity supports the DIVA model prediction that this area contains cell populations representing syllable motor programs without regard for semantic content. Furthermore, the superior paravermal cerebellum is more active for consonant–vowel syllables compared with vowels, perhaps due to increased timing constraints for consonant production.

Keywords

speech motor control; cerebellum; Broca’s area; hemispheric asymmetry; region-of-interest analysis; sparse clustered imaging; syllable production

Speech is one of the most complex tasks carried out by the human brain, involving the integration of auditory information from the temporal lobe, motor information from the frontal lobe, somatosensory information from the parietal lobe, and an array of subcortical structures, including the cerebellum, basal ganglia, and thalamus. A window into how different brain

regions interact to produce a rapid stream of speech sounds can therefore provide insight into a wide array of brain mechanisms. Understanding the brain processes governing the production of the elementary building blocks of speech, such as phonemes and syllables, is a significant step in achieving this goal. However, functional imaging studies of speech production have concentrated primarily on word production and linguistic aspects of speech communication (for a review, see Indefrey & Levelt, 2004) and/or have used syllable production tasks that may involve sequencing processes (e.g., Bohland & Guenther, 2006; Soros et al., 2006). A primary goal of the current study was to investigate, using functional magnetic resonance imaging (fMRI), the premotor lateralization and cerebellar involvement during the articulation of vowels and monosyllables that do not invoke syllable sequencing processes.

Brain activations observed during one of the simplest speech tasks, pseudoword reading (see Table 1), can provide insight into the mechanisms of speech articulation. A meta-analysis of pseudoword reading studies (Indefrey & Levelt, 2004) identified activity in several brain regions. Bilateral sites commonly activated included the posterior inferior frontal gyrus, ventral precentral gyrus, middle superior temporal gyrus, and medial and lateral cerebellum. Left-lateralized activation sites include the supplementary motor area (SMA), posterior fusiform gyrus, middle cingulum, and thalamus. However, several issues limit the information gained from these studies. First, most of them were performed using positron emission tomography (PET) scanners with lower spatial resolution than modern fMRI equipment, thereby impeding precise localization of neural activation. Second, several of these studies involved only covert speech, and clear differences in brain activity between covert and overt speech have been identified (Barch et al., 1999). Third, a number of these studies used sequences of syllables (Bohland & Guenther, 2006; Riecker, Ackermann, Wildgruber, Meyer, et al., 2000) that may invoke syllable sequencing mechanisms. Fourth, a number of these studies involved a repetition paradigm during which participants were asked to repeat the same utterance several times in a row (Riecker, Ackermann, Wildgruber, Meyer, et al., 2000) or to repeat the same utterance on successive trials (Soros et al., 2006). This can reduce the activity due to the repeated stimulus. Finally, the goal of most of these studies was to identify differences in brain activation between word and nonword reading and not the neural processes of speech articulation, which is our primary goal here.

As further motivation for the current study, results from several of the studies listed in Table 1 are contradictory and require further investigation. For example, Riecker, Ackermann, Wildgruber, Meyer, et al. (2000) asked participants to produce consonant–vowel (CV) syllables, a CCCV syllable, and a CVCVCV syllable sequence and reported activity in primary motor, primary sensory, and cerebellar regions. Unexpectedly, they noted a reduction or absence of activation in these regions for the complex utterance “pataka” compared with the simple syllable “ta.” Cerebellar activity was reported only when the syllable “stra” was produced. Lotze, Seggewies, Erb, Grodd, and Birbaumer (2000) investigated brain activity using repetitions of the syllables “pa,” “ta,” “ka,” and the syllable sequence “pataka.” Contrary to the finding by Riecker, Ackermann, Wildgruber, Meyer, et al. (2000), when contrasting the syllable sequence to the single syllables they reported similar activity in primary motor and sensory cortices and greater activity in the SMA for the syllable sequence. Soros et al. (2006) also explored the speech production network using similar stimuli: “ah,” “pa,” “ta,” “ka,” and “pataka.” Similar to Lotze et al. (2000), they reported increased activation when comparing “pataka” production to single-syllable production. However, they reported increases only in superior temporal gyrus and putamen and none in the SMA. Bohland and Guenther (2006) used sequences of CV-pseudowords (e.g., “ta-ki-su,” “stra-stra-stra”) to investigate effects of speech sequencing and syllable complexity. They reported increased activation in the SMA when comparing complex sequences to simple sequences (e.g., “ta-ki-su” vs. “ta-ta-ta”). Among the pseudoword studies that used syllable-level units, Bohland and Guenther (2006) was the only study to report activation in the anterior insula when comparing

syllable sequence production to a visual baseline task. Soros et al. (2006) pointed out that the lack of insular activation in their study could be due to the repetitive nature of their paradigm. These mixed results make it difficult to draw clear conclusions about speech motor control. Differences may be attributed to variations in the task details, scanner strength, scanning protocol, stimuli, and the stimulus presentation methods used and highlight the need for further studies.

Against this background, the goal of the current study was to resolve the previously observed differences by providing an anatomically detailed analysis of the brain regions underlying articulation of the simplest speech sounds using state-of-the-art fMRI. Furthermore, in order to move toward a more detailed functional account of the neural bases of speech production, the current study tested two specific hypotheses concerning the neural mechanisms of speech articulation in different brain regions derived from the DIVA neurocomputational model of speech production (Guenther et al., 2006).

Hypothesis 1: The production of even a single phoneme or simple syllable will result in left-lateralized activation in inferior frontal gyrus, particularly in the pars opercularis region (Brodmann's Area 44; posterior Broca's area) and adjoining ventral premotor cortex.

Although it has long been recognized that control of language production is left lateralized in the inferior frontal cortex (e.g., Broca, 1861; Goodglass 1993; Penfield & Roberts 1959), most speech production tasks show bilateral activation in most areas, including the motor, auditory, and somatosensory cortices (e.g., Bohland & Guenther, 2006; Fiez, Balota, Raichle, & Petersen, 1999). This raises an important question: At what level of the processing hierarchy does language production switch from being largely left lateralized to bilateral in the cerebral cortex?

According to the DIVA model, the transition occurs at the point where the brain transforms syllable and phoneme representations—hypothesized to reside in a *speech sound map* in left inferior frontal cortex (specifically, the ventral premotor cortex and posterior inferior frontal gyrus)—into bilateral motor and sensory cortical activations that control the movements of speech articulators. The model posits that cells in the speech sound map are indifferent to the meaning of the sounds; thus, they should be active even when a speaker is producing elementary nonsemantic utterances. The significance of this hypothesis is that it predicts that storage of the motor programs for speech sounds are left lateralized. These are purely sensorimotor representations with no linguistic meaning. Prior pseudoword studies have failed to note left-lateralized activation in the inferior frontal cortex; here, we directly investigate this issue using statistically powerful region-of-interest (ROI)-based analysis techniques to test the DIVA model's prediction of left-lateralized activity in this region even during production of single nonsense syllables and bisyllables.

Hypothesis 2: The cerebellum—in particular, the superior paravermal region (Guenther et al., 2006; Wildgruber, Ackermann, & Grodd, 2001)—will be more active during CV productions compared with vowel-only productions due to the stricter timing requirements of consonant articulation.

According to the DIVA model, this region of the cerebellum is involved in the encoding of feedforward motor programs for syllable production. Cerebellar damage results in various movement timing disorders, including difficulty with alternating ballistic movements, delays in movement initiation (Inhoff, Diener, Rafal, & Ivry, 1989; Meyer-Lohmann, Hore, & Brooks, 1977), increased movement durations, reduced speed of movement, impaired rhythmic tapping (Ivry, Keele, & Diener, 1988), impaired temporal discrimination of intervals (Nichelli, Alway, & Grafman, 1996; Mangels, Ivry, & Shimizu, 1998), and impaired estimation of the velocity of moving targets (Ivry, 1997). Cerebellar damage can also result in *ataxic dysarthria*, a

disorder characterized by various abnormalities in the timing of motor commands to the speech articulators (Ackermann & Hertrich, 1994; Hirose, 1986; Kent, Kent, Rosenbek, Vorperian, & Weismer, 1997; Kent & Netsell, 1975; Kent, Netsell, & Abbs, 1979; Schonle & Conrad, 1990). The vowel formant structures of people diagnosed with ataxic dysarthria are typically normal, but the transitions to and from consonants are highly variable (Kent, Kent, Duffy, et al., 2000; Kent, Duffy, Slama, Kent, & Clift, 2001). This finding suggests that the cerebellum facilitates the rapid, coordinated movements required for consonant production. If the feedforward commands for syllables are indeed represented in part in the superior paravermal cerebellum as predicted by the DIVA model, then this region should be more active for CV syllables than for single vowels. Again, previous pseudoword studies have not reported such a difference.

To test these hypotheses, we used a sparse, clustered acquisition fMRI technique. Sparse imaging methods (Birn, Bandettini, Cox, & Shaker, 1999; Engelien et al., 2002; Le et al., 2001; MacSweeney et al., 2000; Eden et al., 1999) have been developed to address concerns related to imaging of auditory and speech tasks. These techniques take advantage of the fact that the peak hemodynamic response to a neural event is delayed (Engelien et al., 2002; Amaro et al., 2002), and therefore the response due to an articulatory movement can be measured after the completion of the movement. This allows stimulus presentation and overt responses to occur in relative silence (Amaro et al., 2002), thereby eliminating hemodynamic response to scanner noise that may interfere with auditory cortical activation due to speaking. The sparse acquisition technique also eliminates the susceptibility artifact (Barch et al., 1999) caused by changes in the volume of the air cavity of the mouth during speaking, because these artifacts exist only during the period of movement (Birn, Bandettini, Cox, Jesmanowicz, & Shaker, 1998).

In addition to using an improved scanning paradigm for overt speech tasks, we sought to improve the anatomical specificity of our results by avoiding the methodological shortcomings of standard participant averaging procedures. Group statistical analysis of fMRI data is hindered by (a) noise in the acquired data (Krüger & Glover, 2001) and (b) the striking amount of regional neuroanatomical variability in the location of sulci and gyri across participants, even after transformation into a standard stereotaxic space (e.g., Talairach space; Talairach & Tournoux, 1988). Standard voxel-based fMRI analyses typically address these problems by spatially smoothing data (Lowe & Sorenson, 1997) in three-dimensional (voxel) space. This results in the smearing of functional data across topologically and functionally distant regions. For example, data from auditory cortex in the superior temporal plane are blended with data from motor and/or somatosensory cortices in the opercula located immediately above the temporal plane. Although these areas lie in close proximity in three-dimensional space, they are far apart on the cortical sheet, separated by the insular cortex. Furthermore, a nonlinear normalization procedure, meant to align anatomical regions across participants (Ardekani et al., 2005; Ashburner & Friston, 1999; Woods, Grafton, Watson, Sicotte, & Mazziotta, 1998), leaves a profound amount of residual intersubject anatomical variability (Nieto-Castanon, Ghosh, Tourville, & Guenther, 2003) while further blurring individual anatomical distinctions. Thus, standard voxel-based procedures compromise the anatomical localization of blood-oxygen-level-dependent (BOLD) responses. To overcome this shortcoming, an anatomically precise ROI-based procedure was performed in addition to standard voxel-based analyses. Functional data were pooled within ROIs defined by individual anatomical landmarks prior to performing group statistical analyses. This ROI-based analysis procedure eliminates the need for voxel-based spatial smoothing and nonlinear normalization and has been shown to significantly increase statistical power (Nieto-Castanon et al., 2003).

Method

Participants

Ten right-handed native speakers of American English (3 women, 7 men; 19–47 years of age, $M = 26$ years of age) with no history of neurological disorder participated in the study. All study procedures, including recruitment and acquisition of informed consent, were approved by the institutional review boards of Boston University and Massachusetts General Hospital.

Experimental Protocol

Scanning was performed at the Athinoula A. Martinos Center for Biomedical Imaging in Charlestown, Massachusetts, using Siemens Allegra (6 participants) and Siemens Trio (4 participants) 3T scanners equipped with a whole-head coil. Each trial began with the presentation of a stimulus projected orthographically onto a screen viewable from within the scanner. Three different types of speech stimuli were used: vowels (V), consonant–vowel monosyllables (CV), and CVCV bisyllables. The full stimulus list is provided in Table 2. Nonword stimuli were chosen to minimize brain activation due to semantic processing. A baseline condition during which “xxxx” was projected on screen was also included. Participants were instructed to read each speech stimulus aloud as soon as it appeared on the screen and to remain silent when the control stimulus appeared. Stimuli remained onscreen for 3 s. The participants were asked to pronounce each stimulus in a specific way (they practiced these pronunciations beforehand), and the context of the task, as well as the spelling of the stimuli, discouraged the interpretation of the stimuli as real words.

An experimental run contained 65 trials (13 vowels, 26 CVs, 13 CVCV bisyllables, and 13 baseline trials) that were randomly distributed throughout the run. For each run, stimuli were randomly chosen from the stimulus list in Table 2. Trial length was varied from 15 s to 18 s to prevent the participants from becoming accustomed to a fixed interstimulus interval. Participants completed two to three functional runs, each approximately 20 min long; time constraints and/or participant discomfort prevented some participants from completing the third run. Priming effects associated with repetitive presentation of the same stimulus (Rosen, Buckner, & Dale, 1998) were eliminated by randomizing the order of stimulus conditions. Stimulus presentation was implemented using Psyscope (Macwhinney, Cohen, & Provost, 1997) on an Apple iBook computer. Prior to entering the scanner, participants practiced speaking each stimulus aloud to ensure proper pronunciation. Because vocal output was not recorded in this setup, the participant was monitored intermittently by listening to the audio output of the built-in scanner microphone.

Data Acquisition

BOLD responses resulting from speech production were measured using fMRI (Belliveau et al., 1992; Ogawa, Lee, Kay, & Tank, 1990). Functional data were obtained using a sparse, clustered acquisition, event-related paradigm (Eden, Joseph, Brown, Brown, & Zeffiro, 1999; Engelen et al., 2002; Le, Patel, & Roberts, 2001; MacSweeney et al., 2000). The timeline for a single trial is schematized in Figure 1. Two consecutive whole-head image volumes were acquired beginning 4 s after trial onset. The functional acquisitions were timed to include the peak of the hemodynamic response due to speech production; the peak hemodynamic response has previously been estimated to occur 4–7 s after the onset of articulation (Birn et al., 1998; Engelen et al., 2002). This sparse sampling protocol permitted data acquisition after participants finished speaking—that is, in the absence of noise associated with image collection. This protocol offered several advantages over traditional block and event-related methods: (a) the task was performed in a more natural, quiet-speaking environment, (b) BOLD responses in auditory areas were not compromised by the imaging noise, (c) artifacts due to

head movement and changing oral cavity volumes were avoided, and (d) monitoring participant responses for proper pronunciation of the stimuli was made easier.

Functional volumes were acquired using a single-shot gradient-echo echo planar imaging (EPI) sequence. Each volume consisted of 30 T2*-weighted axial slices (5 mm thick, 0 mm skip) covering the whole brain (time of repetition [TR] = 3 s, time of echo [TE] = 30 ms, flip angle = 90°, field of view [FOV] = 200 mm and interleaved scanning). Slices were oriented parallel to an imaginary plane passing through the anterior and posterior commissures and perpendicular to the sagittal midline. Functional images were reconstructed as a $64 \times 64 \times 30$ matrix with a spatial resolution of $3.1 \times 3.1 \times 5$ mm. To aid in the localization of functional data and for delineating anatomical ROIs, high-resolution T1-weighted anatomical volumes were collected prior to the functional runs with the following parameters: TR = 6.6 ms, TE = 2.9 ms, flip angle = 8°, 128 sagittal slices, 1.33 mm thick, FOV = 256 mm. Anatomical images were reconstructed as a $256 \times 256 \times 128$ matrix with a $1 \times 1 \times 1.33$ mm spatial resolution.

Image Processing

The functional data from each participant were corrected for head movement and coregistered with the high-resolution structural image using the Statistical Parametric Mapping toolbox, version 2 (SPM2; Friston, Ashburner, Holmes, & Poline, 2002). Because of the clustered-sparse scanning procedure used in this study, slice-timing correcting could not be applied. Slice-timing correction relies on collecting image volumes that are temporally close to one another. The coregistered functional data were normalized to Montreal Neurological Institute (MNI) stereotaxic space by two methods: (a) a nonlinear transformation provided the normalized data set used for voxel-based analysis with the SPM2 toolbox and (b) an affine transformation provided the normalized data set used for ROI analysis. After applying each transform, the data were resampled to $2 \times 2 \times 2$ mm resolution. The high-resolution anatomical volume was also normalized according to both the nonlinear and affine normalization procedures. This dual-stream analysis process is schematized in Figure 2. The nonlinear normalized images were smoothed with an isotropic Gaussian kernel with a full width at half maximum (FWHM) of 12 mm.

The affine-normalized anatomical data for each participant was automatically parcellated (Fischl et al., 2004) using Freesurfer (Dale, Fischl, & Sereno, 1999; Fischl, Sereno, & Dale, 1999) into 54 cortical (Caviness, Meyer, Makris, & Kennedy, 1996; Tourville & Guenther, 2003) and 13 subcortical (Fischl et al., 2002) ROIs. The cerebellum was manually parcellated using a graphical interface component of an ROI toolbox (<http://speechlab.bu.edu/SLT.php>) that was developed for interactive parcellation of structural magnetic resonance (MR) images. The automatically parcellated cortical surface ROIs were manually inspected and corrected for gross labeling errors. The parcellated cortical and subcortical ROIs were then converted to a volume mask. The ROI volume masks from each participant were used to extract data from each voxel in that participant's affine-normalized functional data volume. This procedure provides functional ROI data from each participant that reflects individual anatomy.

Functional Data Analysis

Estimates of task-related effects were assessed by applying the general linear model to the normalized functional data. The SPM2 toolbox was used to estimate responses for each voxel based on the nonlinear smoothed data; estimates for each ROI were based on the mean response from each ROI based on the affine-normalized, unsmoothed data using the analysis components of the ROI toolbox. The details of the ROI-based statistical methods are described elsewhere (Nieto-Castanon et al., 2003). There are several key differences between the voxel-based and ROI-based analyses. As noted previously, the most significant difference is that the ROI-based analysis compares responses in anatomically equivalent regions across participants,

eliminating the need to spatially smooth functional data and providing greater statistical sensitivity. The increased detection power afforded by the ROI analysis comes at the loss of spatial specificity provided by voxel-based analyses. In the analysis conducted for this study, we leveraged the advantages offered by both techniques to ensure detection and localization of task-related effects. The voxel-based data and the ROI data were both analyzed using the same design matrix for the general linear model.

The design matrix for the general linear model included one covariate of interest for each test condition. The BOLD response for each trial was modeled using a single-bin finite impulse response (FIR) basis function spanning the time of acquisition of the two consecutive volumes. As a result of the clustered acquisition scheme, global intensity differences between the first and second volume collected during each trial arose due to T1-relaxation related effects. These differences were modeled by a high-frequency regressor ($[1 \ -1 \ 1 \ -1 \ \dots]^T$) that was added to the design matrix. Motion parameters estimated from the realignment process and a linear covariate estimating drift for each run were also added to the design matrix.

Following estimation of the general linear model, mixed effects statistical analysis was performed to test three contrasts of interest. The first compared monosyllable production (V and CV conditions) to the visual baseline condition. The second contrast compared bisyllabic utterances (the CVCV condition) with the monosyllables, and the third contrast compared CV monosyllables with the vowel-only monosyllables (V). Contrast images were determined for each participant by comparing the appropriate condition parameter estimates. Group effects were then assessed by treating participants as random effects and performing one-sample t tests across the individual contrast images.

Unless otherwise noted in the Results section, the results of the standard SPM and ROI analyses were corrected for multiple comparisons by using a statistical threshold that controlled the false discovery rate (FDR; Genovese, Lazar, & Nichols, 2002) at 0.05. The table for results from the standard SPM analyses was generated using the Automatic Anatomical Labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002). For each contrast of interest, the significant voxels were classified and labeled by the atlas, and their corresponding effect sizes were extracted. When multiple significant locations were detected for an AAL label, only the two most active locations were reported.

The ROI data served as input to a nonparametric bootstrapping method (Efron, 1979) designed to assess laterality effects. Specifically, the tests were used to determine whether responses in the ventral premotor, posterior inferior frontal gyrus, and cerebellar ROIs were greater in the left than in the right hemisphere. Details of the bootstrapping method are provided in the Appendix.

Results

Tables 3 and 4 display results from the standard SPM and ROI analyses, respectively.

Monosyllable Production

The standard SPM results for the comparison of monosyllable (V, CV) production to the baseline visual condition are shown in Figure 3. Table 3 reports the effect sizes and MNI coordinates of maximally significant locations of activation from the mixed effects analysis. The FDR correction corresponds to an uncorrected $p < .0056$. Bilaterally activated areas included ventral motor cortex; SMA; pre-supplementary motor area (preSMA); insula; the superior temporal plane (including Heschl's gyrus and planum temporale); posterior superior temporal gyrus; putamen; pallidum; thalamus; and lobules VI, VII (Crus I), VIII, and X and vermis IV/V of the cerebellum. Significant activation in the left hemisphere only was found

in ventral premotor cortex, the inferior frontal gyrus, and the inferior temporal–occipital junction. No areas were active only in the right hemisphere.

The results of the ROI analysis were similar to SPM analysis results (see Figure 4), revealing bilateral activity in frontal operculum, anterior central operculum, posterior central operculum, anterior and posterior insula, preSMA, supplementary motor area, Heschl's gyrus, planum temporale, posterior superior temporal gyrus, anterior medial cerebellum, superior posterior lateral cerebellum, superior posterior medial cerebellum, ventral motor cortex, dorsal motor cortex, ventral somatosensory cortex, putamen, pallidum, and thalamus. Regions that were significantly active only in the left hemisphere were ventral premotor cortex, inferior temporal occipital cortex, inferior frontal gyrus pars opercularis, and, to a lesser extent, pars triangularis. Nonparametric lateralization tests verified significantly greater activation in the left hemisphere for these regions compared to the right hemisphere. The ROI analysis also revealed regions significantly active only in the right hemisphere. These were the planum polare, anterior superior temporal gyrus, and posterior dorsal superior temporal sulcus. The FDR correction corresponds to an uncorrected region level $p < .0228$.

Bisyllables Compared With Monosyllables

The standard SPM mixed effects analysis results comparing bisyllabic utterances with monosyllabic utterances are shown in Figure 5. The FDR correction corresponds to an uncorrected $p < .0022$. Bisyllables showed increased activation bilaterally in preSMA, superior temporal gyrus, lobule VI of the cerebellum, superior parietal cortex, and extra-striate visual cortex. Significant additional left-hemisphere activation was observed only in primary motor cortex, inferior frontal gyrus, and Crus I (located in lobule VII) of the cerebellum. ROI analysis of this contrast (see Table 4) showed bilateral activation in ventral motor cortex, preSMA, Heschl's gyrus, planum temporale, anterior and posterior superior temporal gyrus, posterior dorsal superior temporal sulcus, superior posterior medial cerebellum, putamen, and pallidum. Significant activation found only in the left hemisphere included ventral premotor cortex, inferior frontal gyrus pars opercularis, inferior temporal–occipital junction, superior parietal lobule, and superior posterior lateral cerebellum. The FDR correction for the ROI analysis corresponds to an uncorrected region level $p < .0137$.

CV Monosyllables Compared With Single Vowels

The third contrast, comparing CV monosyllables to vowels, did not show any significant activation when the standard SPM mixed effects analysis results or whole-brain ROI results were corrected for multiple comparisons. However, ROI analysis restricted to the cerebellum revealed significant activation in the right anterior medial cerebellum after controlling the FDR at 0.05. The FDR correction corresponds to an uncorrected region level $p < .003$. Lateralization tests using this ROI did not reveal a statistically significant right-hemisphere lateralization.

Discussion

A sparse imaging technique was used to measure BOLD responses during the production of phonetically simple pseudoword stimuli that do not require syllable sequencing mechanisms. The experiment revealed the minimal network of brain areas involved in motor control of speech. This network includes bilateral ventral somatosensory and motor cortex, superior temporal gyrus (including Heschl's gyrus and the planum temporale), SMA, preSMA, anterior and posterior insula, cerebellum (lobule IV, V, VI, Crus I, and lobule VIII), thalamus, and basal ganglia. In addition, strongly left-lateralized activation was found in the inferior frontal gyrus and ventral premotor cortex. Producing bisyllabic utterances led to additional activity within this basic network, including bilateral motor cortex, preSMA, cerebellar lobule VI, Heschl's gyrus, planum temporale, superior temporal gyrus, left inferior frontal gyrus, ventral premotor

cortex, and Crus I of the cerebellum. Additional activity was found only in left superior parietal lobule and may be related to the processing of additional letters in the orthographic display.

In the following subsections, we discuss the likely roles of these active areas in the speech production process. This discussion will not address the inferior temporal–occipital junction and extra-striate cortex activations because we believe these areas are related to reading of nonsense words rather than speech motor control (Puce, Allison, Asgari, Gore, & McCarthy, 1996).

Posterior Inferior Frontal Gyrus and Ventral Premotor Cortex

The standard SPM and ROI results support Hypothesis 1, showing left lateralized activity in posterior inferior frontal gyrus (particularly the ventral portion of the gyrus, including the frontal operculum) and adjoining ventral premotor cortex during monosyllable production. The DIVA model (Guenther et al., 2006) posits a speech sound map in this region. According to the model, cells in this map represent well-learned speech sounds (syllables and phonemes) without regard for semantic content, and activation of these cells leads to the readout of learned feedforward motor programs for producing the sounds. Activations outside this region were generally bilateral when speech tasks were compared with baseline. Geschwind (1969) concluded that speech planning relied on the left hemisphere, but he could not determine, using his data, at what specific level of the production process speech becomes left lateralized. These results clearly indicate that semantic content is unnecessary for left lateralization in speech; even the production of a single nonsense monosyllable involves left-lateralized circuitry in the inferior frontal cortex. This finding provides an important data point concerning the stage at which speech processing switches from left lateralized to bilateral.

Primary Motor and Somatosensory Cortex

In the current study, the ventral primary motor cortex (BA 4) was the most active ROI in each speech task in both hemispheres. Bilateral activation of this region, which drives movements of the speech articulators, is consistent with all studies of overt word and pseudoword production (see Indefrey & Levelt, 2004, for a review). No significant difference ($p < .05$ at the region level) in this region was noted in the V and CV conditions. The increased ventral motor activity during bisyllabic productions, relative to monosyllabic productions, likely reflects the increased control demands required for longer utterances. Although significant activation was found in both hemispheres, a nonparametric test of laterality demonstrated significantly more activation of ventral motor cortex in the left hemisphere, consistent with the findings of Riecker, Ackermann, Wildgruber, Dogil, & Grodd (2000) and Bohland and Guenther (2006).

Significant activity was also noted in a portion of dorsal primary motor cortex bilaterally in all speech tasks. Activation of this area was not modulated by increasing the number of syllables. This region may correspond to the dorsal-most speech articulator motor representation or may subserve respiratory control, which requires the maintenance of subglottal pressure during speech.

Bilateral activation of ventral primary somatosensory cortex (BA 1,2,3) was found for each speaking condition relative to baseline, and the lateralization test did not find any differences in activity between the two hemispheres. Activation in this area is expected during overt speech, which involves the processing of tactile and proprioceptive feedback from vocal tract articulators. The somatosensory representation of the speech articulators along the ventral postcentral gyrus has been shown to mirror the precentral motor representations (see Guenther et al., 2006, for a review).

SMA and preSMA

The SMA and preSMA have been implicated in the initiation and/or sequencing of motor acts in a number of neuroimaging and electrophysiological experiments. Electrical stimulation mapping of the somatotopic organization of SMA has shown the head and face areas to be represented in the anterior region (Fried et al., 1991); vocalization and speech arrest or slowing of speech are evoked with electrical stimulation to the preSMA anterior to the SMA face representation.

In the current study, SMA and preSMA activations were found for each speaking condition compared to baseline, and the laterality test did not reveal any significant difference between the hemispheres. There was also greater activity bilaterally in preSMA when the bisyllabic condition was compared with the monosyllabic conditions, similar to the findings of Lotze et al. (2000). The latter result makes sense when one considers the increased sequencing load for a bisyllabic utterance compared with a monosyllable. In a recent study of syllable sequence production, Bohland and Guenther (2006) found that increased complexity of a syllable sequence (e.g., “ta-ki-su” vs. “ta-ta-ta”) led to greater activity in preSMA, as did increased complexity of the syllables within a sequence (e.g., “stra” vs. “ta”). However, both Riecker, Ackermann, Wildgruber, Meyer, et al. (2000) and Soros et al. (2006) reported no increase in the SMA or preSMA as a consequence of increased complexity of the utterance. These studies used a task paradigm in which each of the task conditions was repeated several times in a sequence. The preSMA has been implicated in preparatory activities related to movement execution (Lee, Chang, & Roh, 1999). Therefore, repetitive tasks that require preparation only on the first trial may show no differences in this region between the different task conditions.

Based on their fMRI study, Alario, Chainay, Lehericy, & Cohen (2006) suggest that preSMA is used for word selection, whereas SMA is primarily used for motor execution. They also report a subdivision within the preSMA, with the anterior part more involved with lexical selection and the posterior part involved with sequence encoding and execution. It is in the more posterior “sequence encoding” region of the preSMA where increased activation with increased syllabic content is seen in the current study, in keeping with the Alario et al. (2006) proposal. In summary, the current study adds further support to the view that SMA is more involved in the execution of speech output, whereas the preSMA is more involved in the process of sequencing syllables and phonemes.

Cerebellum

Comparing production of nonsense monosyllables to the baseline visual condition revealed activations in three distinct regions of the cerebellar cortex: (a) a superior, anterior vermal/paravermal region (lobules IV and V), (b) a region in lobule VI and Crus I bilaterally that lies behind the primary fissure in the intermediate portion of the cerebellar hemisphere (lateral and posterior to the lobule IV and V activations), and (c) a region in the anterior portion of lobule VIII bilaterally near the border between the intermediate and lateral portions of the inferior cerebellar hemisphere. The superior vermal/paravermal cerebellar cortex has been implicated in previous lesion studies of ataxic dysarthria (e.g., Ackermann, Vogel, Petersen, & Poremba, 1992; Urban et al., 2003).

The ROI results also support Hypothesis 2—that the cerebellum should be more active for CV productions compared with vowel productions. The results of the contrast comparing CV utterances with vowel productions showed greater activity in the anterior medial portion of the right superior cerebellar hemisphere; no increases were found in the cerebral cortex. According to the DIVA model (Guenther et al., 2006), this region of the cerebellum encodes feedforward motor commands for producing syllables and phonemes. The model is thus compatible with

the finding of more cerebellar activity in the CV condition compared with the vowel (V) condition.

Bohland and Guenther (2006) reported increased activity in right inferior cerebellum in response to increased complexity of sequences. In this study, we did not observe an increase in this region when comparing the bisyllable utterances to the monosyllable utterances; the only increase was observed in the superior cerebellar cortex. The task in the Bohland and Guenther (2006) study involved a delayed response and therefore a working memory component. Other verbal working memory tasks (Chen & Desmond, 2005; Kirschen, Chen, Schraedley-Desmond, & Desmond, 2005) have also reported activity in the right inferior cerebellum. The lack of a working memory component in the present study may account for the lack of an increase in right inferior cerebellar activity for bisyllables compared with monosyllables.

Primary and Higher Order Auditory Cortices

Activation of auditory cortex during speech is widely thought to reflect hearing one's own voice during speech and possibly is compared to auditory goals for the speech sounds being produced (Guenther et al., 2006). Higher-order auditory regions such as the dorsal bank of the posterior superior temporal sulcus have been shown to be selective to voice, compared with environmental and nonvocal sounds (Belin, Zatorre, Lafaille, Ahad, & Pike, 2000). Regions that are active during both the perception and production of speech include planum temporale, posterior superior temporal gyrus, and posterior superior temporal sulcus (e.g., Buchsbaum, Hickok, & Humphries, 2001; Hickok & Poeppel, 2000); these regions have been hypothesized to include auditory error cells that code the difference between the expected auditory signal and the actual sensory input for feedback control of speech (Guenther et al., 2006).

We found bilateral activation of a wide range of auditory cortical areas when comparing monosyllable production to the visual baseline. These areas included the primary auditory cortex (Heschl's gyrus) as well as higher-order auditory cortical areas. Right-lateralized activity found in planum polare and anterior superior temporal sulcus may reflect processing of pitch information (Zatorre, Evans, Meyer, & Gjedde, 1992), whereas the activity in posterior dorsal superior temporal sulcus may be related to monitoring of auditory feedback (Tourville, Reilly, & Guenther, 2008). No differences were noted in these regions across the V and CV conditions. Additional activity in these regions when comparing bisyllables to monosyllables likely reflects the longer sound duration. Overall, these results are consistent with the available literature while providing more anatomical specificity concerning the locations of auditory activations during speech production.

Insula

In the current study, we observed bilateral activation of both the anterior and posterior insula during monosyllable production. There were no significant differences in insular activity among the three speaking conditions. Although the bulk of the literature concerning insula and speech production focuses on left anterior insula, bilateral activation has been found during simple speech tasks in both anterior (e.g., Bohland & Guenther, 2006) and posterior insula (e.g., Soros et al., 2006).

Soros et al. (2006), Lotze et al. (2000), and Riecker, Ackermann, Wildgruber, Meyer, et al. (2000) did not observe activation in the anterior insula when speaking was compared to their respective baseline tasks. Soros et al. (2006) suggest that the repetitive nature of their task may be responsible for the lack of insula activation, whereas studies that involve more stimulus variation such as the current study show anterior insular activation (e.g., Bohland & Guenther 2006; Wise, Greene, Buchel, & Scott, 1999). This suggestion is supported by Nota and Honda

(2003), who demonstrated activation of the anterior insula when participants were prompted to produce a series of nonrepeating syllable sequences but not when the same sequence was repeated continually. The lack of anterior insula activation during tasks that instruct participants to repeat the particular syllable with metronome-like timing (Lotze et al., 2000; Riecker, Ackermann, Wildgruber, Meyer, et al., 2000) thus appears to be related to the repetitive nature of the tasks. On the basis of lesion studies (Dronkers, 1996; Hillis et al., 2004) and fMRI studies (Bohland & Guenther, 2006), it has been suggested that this region may play a role in sequencing or initiation of speech. The current study paradigm cannot explicitly test these claims, and it is difficult to assign functional contributions to insular cortex activations found in the current study at this time.

Basal Ganglia and Thalamus

Although the precise functional contributions of the basal ganglia and thalamus in speech motor control are still debated, it is well-established that damage to these regions is associated with several disturbances in speech. Damage to the putamen and caudate of the basal ganglia can result in inaccuracies in articulation (Pickett, Kuniholm, Protopapas, Friedman, & Lieberman, 1998)—for example, during a switch from one articulatory gesture to another. The basal ganglia and thalamus form components of multiple parallel cortico-striato-thalamo-cortical loops (Alexander, Crutcher, & DeLong, 1990; Alexander, DeLong, & Strick, 1986). One such loop, known to be involved with motor control, includes the SMA, the putamen, the pallidum, and the ventrolateral thalamus (Alexander et al., 1986). Electrical stimulation studies (for a review, see Johnson & Ojemann, 2000) have suggested involvement of the left ventrolateral thalamus in the motor control of speech, including respiration. The present study shows this entire loop to be activated for the production of monosyllables compared to baseline. The comparison of bisyllabic utterances to monosyllables also showed bilateral increases in activation in putamen and pallidum but did not show any significant increase in activation in the thalamus.

Computational models (Redgrave, Prescott, & Gurney, 1999; Prescott, González, Gurney, Humphries, & Redgrave, 2006) have suggested that the functional role of the basal ganglia is to select and enable the execution of a particular action from a set of competing alternatives. This is particularly relevant during speech production, which involves selecting and sequencing appropriate smaller units (phonemes, syllables) to form larger constructs (words, phrases, sentences). The activation of the “motor loop” of Alexander et al. (1986) in the present study involving extremely simple speech utterances indicates that this circuit is used at relatively low levels in speech motor control.

Methodological Concerns

Although the model-based hypotheses were successfully tested, there are possible methodological issues that can be addressed better in future work. First, despite careful attempts to remove any semantic component to our task, there is always the possibility that a nonsense syllable stimulus will invoke some semantic activity in the brain, considering that all syllables occur in words. Second, sparse clustered acquisition coupled with the lack of audio recordings necessitated the use of a FIR basis function for the data analyses. In future work, participant-specific hemodynamic response functions may be used to generate a better model.

Summary

The results of the experiment presented here demonstrate that a large network of cortical and subcortical regions interacts to produce even the simplest verbal output. This experiment adds detailed data to the scant imaging literature concerning the production of phonetically simple sounds. The results showed that simple pseudoword utterances, in the form of vowels and

monosyllables, generate left-lateralized responses in premotor cortex and posterior inferior frontal gyrus, with bilateral activation in other areas involved in speech motor control. This supports the DIVA model prediction that a speech sound map in the left inferior frontal region represents the lowest level of the speech production process, which is highly left lateralized. Contrasting bisyllables with monosyllables revealed increased activity in several areas that were part of the same network used for monosyllables. Finally, additional activation in the cerebellum was observed when comparing CV syllables to vowels. This may reflect the stricter timing requirements for consonant utterances compared with vowel utterances. These findings are consistent with several neuroanatomical/functional proposals embodied by the DIVA model of speech production (Guenther et al., 2006).

Acknowledgements

This work was supported by NIH Grants R01 DC02852 (awarded to Frank Guenther, PI) and R01 DC01925 (awarded to Joseph Perkell, PI). Use of the Athinoula A. Martinos Center for Biomedical Imaging was aided by support from the National Center for Research Resources Grant P41RR14075 and the MIND Institute. We thank Jason Bohland, Bruce Fischl, Mary Foley, Julie Goodman, Alfonso Nieto-Castanon, Lawrence Wald, and Lawrence White for their help with various aspects of the project and the Athinoula A. Martinos Center for Biomedical Imaging at Massachusetts General Hospital for the use of their facilities.

References

- Ackermann H, Hertrich I. Speech rate and rhythm in cerebellar dysarthria: An acoustic analysis of syllabic timing. *Folia Phoniatrica et Logopaedica* 1994;46:70–78. [PubMed: 8173615]
- Ackermann H, Hertrich I. Voice onset time in ataxic dysarthria. *Brain and Language* 1997;56:321–333. [PubMed: 9070415]
- Ackermann H, Vogel M, Petersen D, Poremba M. Speech deficits in ischaemic cerebellar lesions. *Journal of Neurology* 1992;239:223–227. [PubMed: 1597689]
- Alario FX, Chainay H, Lehericy S, Cohen L. The role of the supplementary motor area (SMA) in word production. *Brain Research* 2006;1076:129–143. [PubMed: 16480694]
- Alexander GE, Crutcher MD, DeLong MR. Basal ganglia–thalamocortical circuits: Parallel substrates for motor, oculomotor, ‘prefrontal’ and ‘limbic’ functions. *Progress in Brain Research* 1990;85:119–146. [PubMed: 2094891]
- Alexander GE, DeLong MR, Strick PL. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience* 1986;9:357–381.
- Amaro E Jr, Williams SC, Shergill SS, Fu CH, MacSweeney M, Picchioni MM, et al. Acoustic noise and functional magnetic resonance imaging: Current strategies and future prospects. *Journal of Magnetic Resonance Imaging* 2002;16:497–510. [PubMed: 12412026]
- Ardekani BA, Guckemus S, Bachman A, Hoptman MJ, Wojtaszek M, Nierenberg J. Quantitative comparison of algorithms for inter-subject registration of 3D volumetric brain MRI scans. *Journal of Neuroscience Methods* 2005;142:67–76. [PubMed: 15652618]
- Ashburner J, Friston KJ. Nonlinear spatial normalization using basis functions. *Human Brain Mapping* 1999;7:254–266. [PubMed: 10408769]
- Barch DM, Sabb FW, Carter CS, Braver TS, Noll DC, Cohen JD. Overt verbal responding during fMRI scanning: Empirical investigations of problems and potential solutions. *Neuroimage* 1999;10:642–657. [PubMed: 10600410]
- Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B. Voice-selective areas in human auditory cortex. *Nature* 2000 January 20;403:309–312. [PubMed: 10659849]
- Belliveau JW, Kwong KK, Kennedy DN, Baker JR, Stern CE, Benson R, et al. Magnetic resonance imaging mapping of brain function: Human visual cortex. *Investigative Radiology* 1992;27(Suppl 2):S59–S65. [PubMed: 1468876]
- Birn RM, Bandettini PA, Cox RW, Jesmanowicz A, Shaker R. Magnetic field changes in the human brain due to swallowing or speaking. *Magnetic Resonance in Medicine* 1998;40:55–60. [PubMed: 9660553]

- Birn RM, Bandettini PA, Cox RW, Shaker R. Event-related fMRI of tasks involving brief motion. *Human Brain Mapping* 1999;7:106–114. [PubMed: 9950068]
- Bohland JW, Guenther FH. An fMRI investigation of syllable sequence production. *Neuroimage* 2006;32:821–841. [PubMed: 16730195]
- Broca P. Remarques sur le siège de la faculté de langage articulé, suivie d'une observation d'aphémie (perte de la parole) [Remarks on the seat of the faculty of articulate language, followed by an observation of aphemia]. *Bulletins de la Société Anatomique de Paris* 1861;6:330–357.
- Brunswick N, McCrory E, Price CJ, Frith CD, Frith U. Explicit and implicit processing of words and pseudowords by adult developmental dyslexics: A search for Wernicke's Wortschatz? *Brain* 1999;122(Pt 10):1901–1917. [PubMed: 10506092]
- Buchsbaum BR, Hickok G, Humphries C. Role of left posterior superior temporal gyrus in phonological processing for speech perception and production. *Cognitive Science* 2001;25:663–678.
- Caviness VS, Meyer J, Makris N, Kennedy DN. MRI-based topographic parcellation of human neocortex: An anatomically specified method with estimate of reliability. *Journal of Cognitive Neuroscience* 1996;8:566–587.
- Chen SHA, Desmond JE. Cerebrocerebellar networks during articulatory rehearsal and verbal working memory tasks. *Neuroimage* 2005;24:332–338. [PubMed: 15627576]
- Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. I: Segmentation and surface reconstruction. *Neuroimage* 1999;9:179–194. [PubMed: 9931268]
- Dronkers NF. A new brain region for coordinating speech articulation. *Nature* 1996 November 14;384:159–161. [PubMed: 8906789]
- Eden GF, Joseph JE, Brown HE, Brown CP, Zeffiro TA. Utilizing hemodynamic delay and dispersion to detect fMRI signal change without auditory interference: The behavior interleaved gradients technique. *Magnetic Resonance in Medicine* 1999;41:13–20. [PubMed: 10025606]
- Efron B. Bootstrap methods: Another look at jackknife. *Annals of Statistics* 1979;7:1–26.
- Engelien A, Yang Y, Engelien W, Zonana J, Stern E, Silbersweig DA. Physiological mapping of human auditory cortices with a silent event-related fMRI technique. *Neuroimage* 2002;16:944–953. [PubMed: 12202082]
- Fiez JA, Balota DA, Raichle ME, Petersen SE. Effects of lexicality, frequency, and spelling-to-sound consistency on the functional anatomy of reading. *Neuron* 1999;24:205–218. [PubMed: 10677038]
- Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, et al. Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron* 2002;33:341–355. [PubMed: 11832223]
- Fischl B, Sereno MI, Dale AM. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage* 1999;9:195–207. [PubMed: 9931269]
- Fischl B, van der Kouwe A, Destrieux C, Halgren E, Segonne F, Salat DH, et al. Automatically parcellating the human cerebral cortex. *Cerebral Cortex* 2004;14:11–22. [PubMed: 14654453]
- Fried I, Katz A, McCarthy G, Sass KJ, Williamson P, Spencer SS, Spencer DD. Functional organization of human supplementary motor cortex studied by electrical stimulation. *Journal of Neuroscience* 1991;11:3656–3666. [PubMed: 1941101]
- Friston, KJ.; Ashburner, J.; Holmes, A.; Poline, J. Statistical parametric mapping. 2002. Available from <http://www.fil.ion.ucl.ac.uk/spm>
- Genovese CR, Lazar NA, Nichols T. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* 2002;15:870–878. [PubMed: 11906227]
- Geschwind, N. Problems in the anatomical understanding of the aphasias. In: Benton, AL., editor. *Contributions to clinical neuropsychology*. Chicago: Aldine; 1969. p. 107-128.
- Goodglass, H. *Understanding aphasia*. New York: Academic Press; 1993.
- Guenther FH, Ghosh SS, Tourville JA. Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language* 2006;96:280–301. [PubMed: 16040108]
- Hagoort P, Indefrey P, Brown C, Herzog H, Steinmetz H, Seitz RJ. The neural circuitry involved in the reading of German words and pseudowords: A PET study. *Journal of Cognitive Neuroscience* 1999;11:383–398. [PubMed: 10471847]

- Herbster AN, Mintun MA, Nebes RD, Becker JT. Regional cerebral blood flow during word and nonword reading. *Human Brain Mapping* 1997;5:84–92. [PubMed: 10096413]
- Hickok G, Poeppel D. Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences* 2000;4:131–138. [PubMed: 10740277]
- Hillis AE, Work M, Barker PB, Jacobs MA, Breese EL, Maurer K. Re-examining the brain regions crucial for orchestrating speech articulation. *Brain* 2004;127(Pt 7):1479–1487. [PubMed: 15090478]
- Hirose H. Pathophysiology of motor speech disorders (dysarthria). *Folia Phoniatica (Basel)* 1986;38:61–88.
- Indefrey P, Levelt WJM. The spatial and temporal signatures of word production components. *Cognition* 2004;92:101–144. [PubMed: 15037128]
- Inhoff AW, Diener HC, Rafal RD, Ivry R. The role of cerebellar structures in the execution of serial movements. *Brain* 1989;112(Pt 3):565–581. [PubMed: 2731021]
- Ivry R. Cerebellar timing systems. *International Review of Neurobiology* 1997;41:555–573. [PubMed: 9378608]
- Ivry RB, Keele SW, Diener HC. Dissociation of the lateral and medial cerebellum in movement timing and movement execution. *Experimental Brain Research* 1988;73:167–180.
- Johnson MD, Ojemann GA. The role of the human thalamus in language and memory: Evidence from electrophysiological studies. *Brain and Cognition* 2000;42:218–230. [PubMed: 10744921]
- Kent R, Netsell R. A case study of an ataxic dysarthric: Cineradiographic and spectrographic observations. *Journal of Speech and Hearing Disorders* 1975;40:115–134. [PubMed: 1123918]
- Kent RD, Duffy JR, Slama A, Kent JF, Clift A. Clinicoanatomic studies in dysarthria: Review, critique, and directions for research. *Journal of Speech, Language and Hearing Research* 2001;44:535–551.
- Kent RD, Kent JF, Duffy JR, Thomas JE, Weismer G, Stuntebeck S. Ataxic dysarthria. *Journal of Speech, Language, and Hearing Research* 2000;43:1275–1289.
- Kent RD, Kent JF, Rosenbek JC, Vorperian HK, Weismer G. A speaking task analysis of the dysarthria in cerebellar disease. *Folia Phoniatica et Logopaedica* 1997;49:63–82. [PubMed: 9197089]
- Kent RD, Netsell R, Abbs JH. Acoustic characteristics of dysarthria associated with cerebellar disease. *Journal of Speech and Hearing Research* 1979;22:627–648. [PubMed: 502519]
- Kirschen MP, Chen SHA, Schraedley-Desmond P, Desmond JE. Load- and practice-dependent increases in cerebro-cerebellar activation in verbal working memory: An fMRI study. *Neuroimage* 2005;24:462–472. [PubMed: 15627588]
- Krüger G, Glover GH. Physiological noise in oxygenation-sensitive magnetic resonance imaging. *Magnetic Resonance in Medicine* 2001;46:631–637. [PubMed: 11590638]
- Le TH, Patel S, Roberts TP. Functional MRI of human auditory cortex using block and event-related designs. *Magnetic Resonance in Medicine* 2001;45:254–260. [PubMed: 11180433]
- Lee KM, Chang KH, Roh JK. Subregions within the supplementary motor area activated at different stages of movement preparation and execution. *Neuroimage* 1999;9:117–123. [PubMed: 9918733]
- Lotze M, Seggewies G, Erb M, Grodd W, Birbaumer N. The representation of articulation in the primary sensorimotor cortex. *Neuroreport* 2000;11:2985–2989. [PubMed: 11006980]
- Lowe MJ, Sorenson JA. Spatially filtering functional magnetic resonance imaging data. *Magnetic Resonance in Medicine* 1997;37:723–729. [PubMed: 9126946]
- MacSweeney M, Amaro E, Calvert GA, Campbell R, David AS, McGuire P, et al. Silent speech-reading in the absence of scanner noise: An event-related fMRI study. *Neuroreport* 2000;11:1729–1733. [PubMed: 10852233]
- Macwhinney B, Cohen J, Provost J. The psyscope experiment-building system. *Spatial Vision* 1997;11:99–101. [PubMed: 9304758]
- Mangels JA, Ivry RB, Shimizu N. Dissociable contributions of the prefrontal and neocerebellar cortex to time perception. *Brain Research: Cognitive Brain Research* 1998;7(1):15–39. [PubMed: 9714713]
- Mechelli A, Friston KJ, Price CJ. The effects of presentation rate during word and pseudoword reading: A comparison of PET and fMRI. *Journal of Cognitive Neuroscience* 2000;12(Suppl 2):145–156. [PubMed: 11506654]

- Mechelli A, Gorno-Tempini ML, Price CJ. Neuroimaging studies of word and pseudoword reading: Consistencies, inconsistencies, and limitations. *Journal of Cognitive Neuroscience* 2003;15:260–271. [PubMed: 12676063]
- Meyer-Lohmann J, Hore J, Brooks VB. Cerebellar participation in generation of prompt arm movements. *Journal of Neurophysiology* 1977;40:1038–1050. [PubMed: 409807]
- Nichelli P, Alway D, Grafman J. Perceptual timing in cerebellar degeneration. *Neuropsychologia* 1996;34:863–871. [PubMed: 8822733]
- Nieto-Castanon A, Ghosh SS, Tourville JA, Guenther FH. Region of interest based analysis of functional imaging data. *Neuroimage* 2003;19:1303–1316. [PubMed: 12948689]
- Nota, Y.; Honda, K. Possible role of the anterior insula in articulation. In: Palethorpe, S.; Tabain, M., editors. *Proceedings of the Sixth International Seminar on Speech Production*. Sydney, Australia: 2003 Dec. p. 191-194.
- Ogawa S, Lee TM, Kay AR, Tank DW. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences of the United States of America* 1990;87:9868–9872. [PubMed: 2124706]
- Penfield, W.; Roberts, L. *Speech and brain-mechanisms*. Princeton, NJ: Princeton University Press; 1959.
- Petersen SE, Fox PT, Snyder AZ, Raichle ME. Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli. *Science* 1990 August 31;249:1041–1044. [PubMed: 2396097]
- Pickett ER, Kuniholm E, Protopapas A, Friedman J, Lieberman P. Selective speech motor, syntax and cognitive deficits associated with bilateral damage to the putamen and the head of the caudate nucleus: A case study. *Neuropsychologia* 1998;36:173–188. [PubMed: 9539237]
- Prescott TJ, González FMM, Gurney K, Humphries MD, Redgrave P. A robot model of the basal ganglia: Behavior and intrinsic processing. *Neural Networks* 2006;19:31–61. [PubMed: 16153803]
- Puce A, Allison T, Asgari M, Gore JC, McCarthy G. Differential sensitivity of human visual cortex to faces, letterstrings, and textures: A functional magnetic resonance imaging study. *Journal of Neuroscience* 1996;16:5205–5215. [PubMed: 8756449]
- Redgrave P, Prescott TJ, Gurney K. The basal ganglia: A vertebrate solution to the selection problem? *Neuroscience* 1999;89:1009–1023. [PubMed: 10362291]
- Riecker A, Ackermann H, Wildgruber D, Dogil G, Grodd W. Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula, and cerebellum. *Neuroreport* 2000;11:1997–2000. [PubMed: 10884059]
- Riecker A, Ackermann H, Wildgruber D, Meyer J, Dogil G, Haider H, Grodd W. Articulatory/phonetic sequencing at the level of the anterior perisylvian cortex: A functional magnetic resonance imaging (fMRI) study. *Brain and Language* 2000;75:259–276. [PubMed: 11049668]
- Rosen BR, Buckner RL, Dale AM. Event-related functional MRI: Past, present, and future. *Proceedings of the National Academy of Sciences of the United States of America* 1998;95:773–780. [PubMed: 9448240]
- Rumsey JM, Horwitz B, Donohue BC, Nace K, Maisog JM, Andreason P. Phonological and orthographic components of word recognition. A PET-rCBF study. *Brain* 1997;120(Pt 5):739–759. [PubMed: 9183247]
- Schonle PW, Conrad B. Differential effects of automatic versus spontaneous speech on peak velocity of inspiratory movements. *Folia Phoniatica (Basel)* 1990;42:239–244.
- Soros P, Sokoloff LG, Bose A, McIntosh AR, Graham SJ, Stuss DT. Clustered functional MRI of overt speech production. *Neuroimage* 2006;32:376–387. [PubMed: 16631384]
- Tagamets MA, Novick JM, Chalmers ML, Friedman RB. A parametric approach to orthographic processing in the brain: An fMRI study. *Journal of Cognitive Neuroscience* 2000;12:281–297. [PubMed: 10771412]
- Talairach, J.; Tournoux, P. *Co-planar stereotaxic atlas of the human brain*. Stuttgart, Germany: G. Thieme; 1988.
- Tourville, JA.; Guenther, FH. *A cortical and cerebellar parcellation system for speech studies*. Boston, MA: Boston University; 2003. [Technical report CAS/CNS-03-022]
- Tourville JA, Reilly KJ, Guenther FH. Neural mechanisms underlying auditory feedback control of speech. *NeuroImage* 2008;39:1429–1443. [PubMed: 18035557]

- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 2002;15:273–289. [PubMed: 11771995]
- Urban PP, Marx J, Hunsche S, Gawehn J, Vucurevic G, Wicht S, et al. Cerebellar speech representation: Lesion topography in dysarthria as derived from cerebellar ischemia and functional magnetic resonance imaging. *Archives of Neurology* 2003;60:965–972. [PubMed: 12873853]
- Wildgruber D, Ackermann H, Grodd W. Differential contributions of motor cortex, basal ganglia, and cerebellum to speech motor control: Effects of syllable repetition rate evaluated by fMRI. *Neuroimage* 2001;13:101–109. [PubMed: 11133313]
- Wise RJ, Greene J, Buchel C, Scott SK. Brain regions involved in articulation. *Lancet* 1999;353:1057–1061. [PubMed: 10199354]
- Woods RP, Grafton ST, Watson JD, Sicotte NL, Mazziotta JC. Automated image registration. II: Intersubject validation of linear and nonlinear models. *Journal of Computer Assisted Tomography* 1998;22:153–165. [PubMed: 9448780]
- Zatorre RJ, Evans AC, Meyer E, Gjedde A. Lateralization of phonetic and pitch discrimination in speech processing. *Science* 1992 May;256:846–849. [PubMed: 1589767]

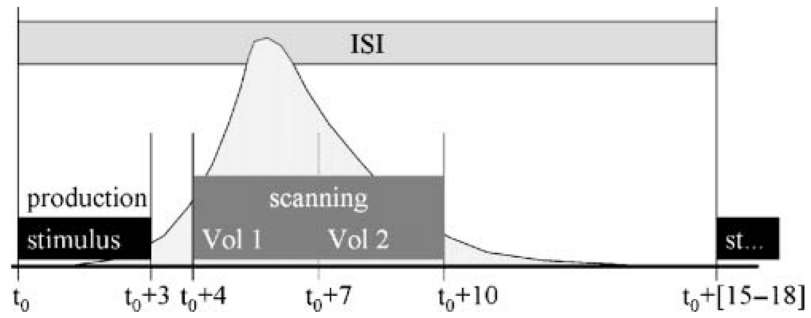


Figure 1.

Event-triggered paradigm used for this study. The stimulus presentation and fMRI volume acquisition timeline are shown in the figure. Each trial consisted of presenting the stimulus, waiting for a short period of time for the hemodynamic response (schematized by the light shaded curve) to increase and then acquiring two whole-brain image volumes (Vol 1, Vol 2) over a 6-s time period beginning 4 s after stimulus onset. This range was chosen to cover the peak of the response.

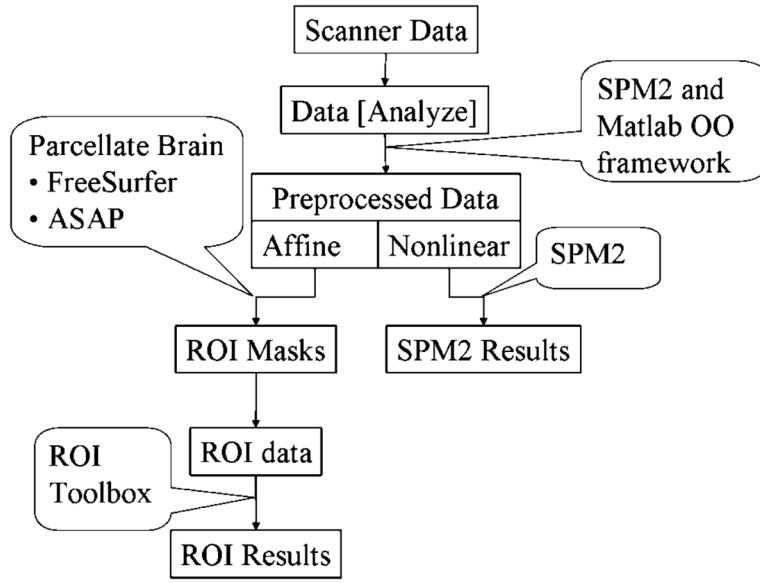


Figure 2. Data processing stream for fMRI data analysis. The data acquired from the scanner are converted to Analyze format and preprocessed (realignment, coregistration, normalization and smoothing) using the SPM toolbox. Two pipelines are started in the normalization process, one involving full (nonlinear) normalization and another involving affine normalization. The full normalized images are smoothed using a 12-mm FWHM Gaussian kernel and then used for fixed effects and mixed effects analysis using the SPM toolbox. The affine normalized images are masked with ROI masks generated using ASAP (A Segmentation and Parcellation Toolbox for Matlab) and FreeSurfer and analyzed separately using an ROI analysis toolbox. The results from each stream are then compared to ensure consistency between the two methods for the major contrasts.

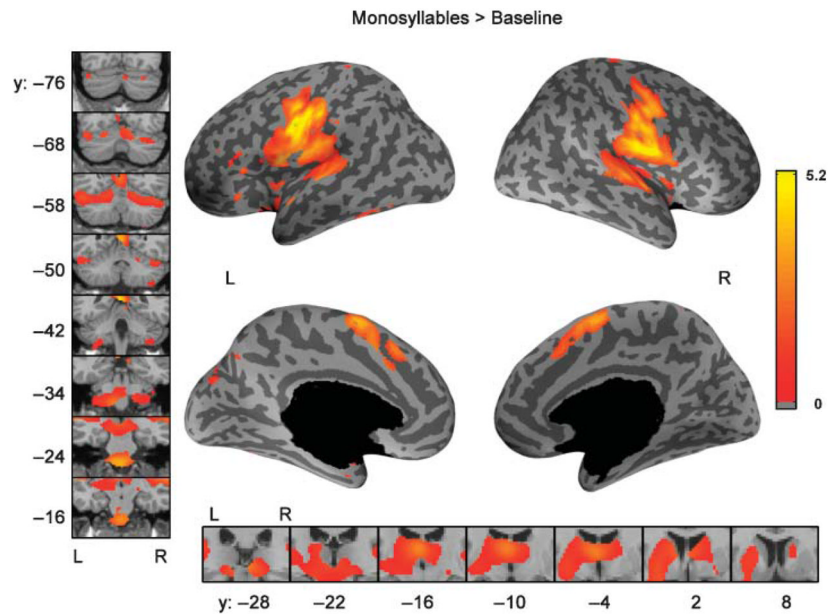


Figure 3.

Comparison of speaking a single monosyllable to the baseline visual condition, standard SPM results. Color indicates the effect size (% signal change of the global mean) from the mixed effects analysis using the SPM toolbox corrected for multiple comparisons using an FDR of 0.05. On the inflated cortical surfaces, the dark gray corresponds to sulci and the light gray corresponds to gyri, while black represents non-cortical regions. The left panel shows coronal slices covering the extent of activation in the cerebellum while the bottom panel shows coronal slices covering the extent of activation in the basal ganglia and thalamus. Y-values refer to planes in MNI-space. The color scale is common to all renderings.

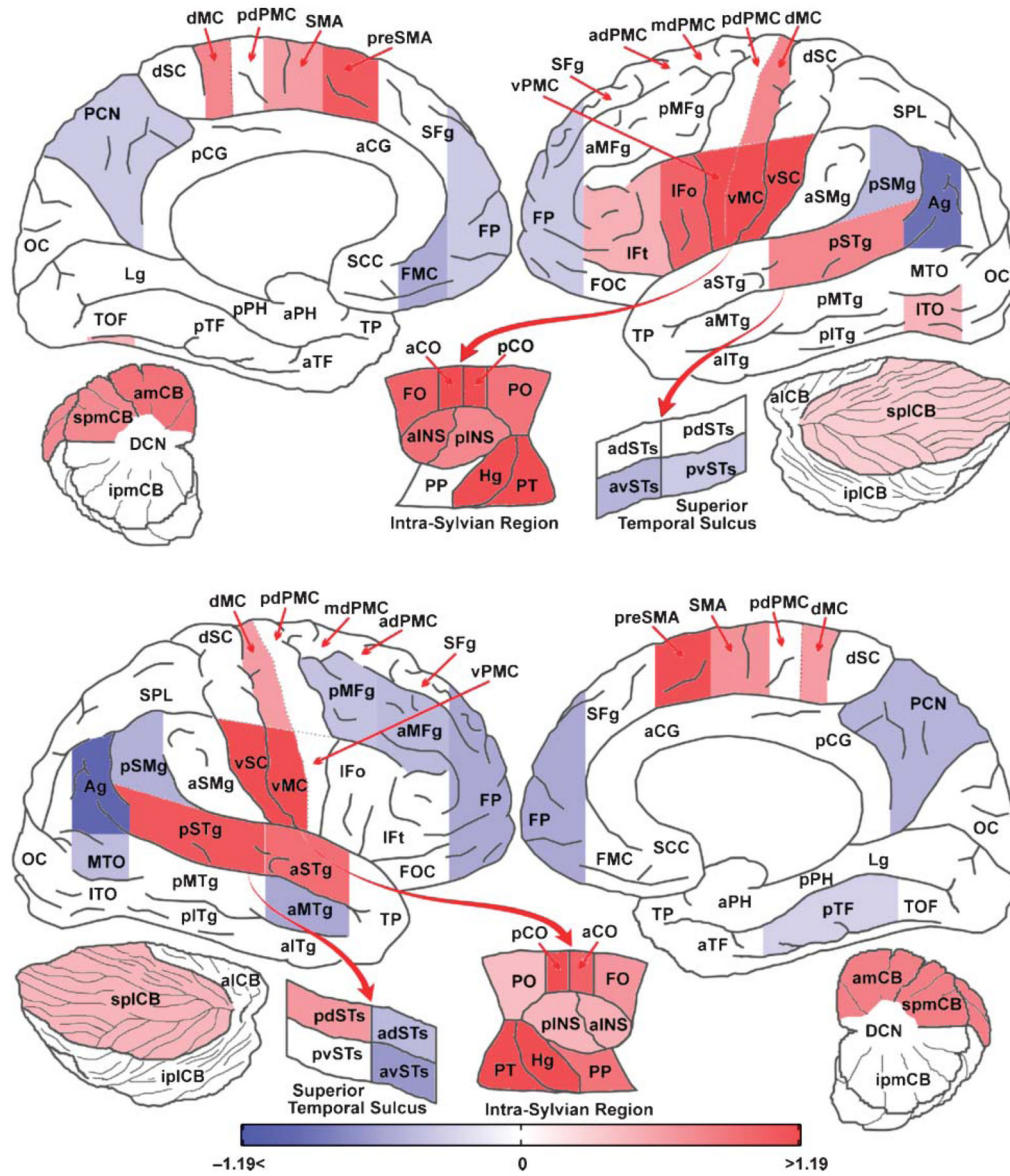


Figure 4. Comparison of speaking a single monosyllable to the baseline visual condition, ROI analysis. Color indicates the effect size (% signal change of the global mean) from the mixed effects ROI analysis corrected for multiple comparisons using an FDR of 0.05. Areas in red are significantly more active in the monosyllable condition and those in blue are more active in the baseline visual condition. See the caption accompanying Table 4 for a description of the ROI abbreviations.

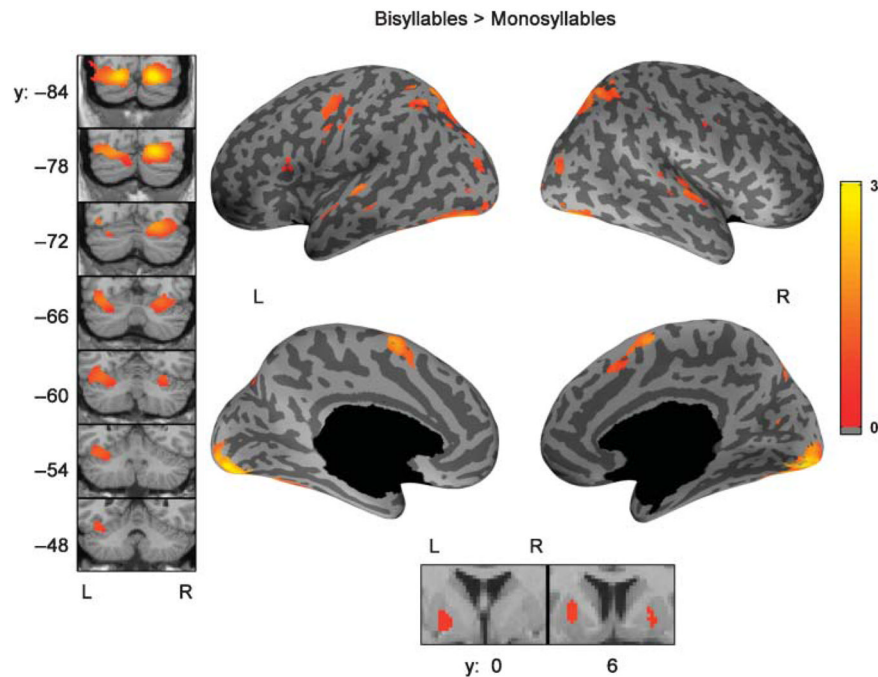


Figure 5. Comparison of bisyllabic utterances to monosyllabic utterances, standard SPM analysis. The regions showing increased activity bilaterally for bisyllables are SMA, the superior cerebellum, superior parietal cortex, extra-striate cortex and superior temporal gyrus. Significant activity in the inferior frontal gyrus and motor cortex was found only in the left hemisphere. The data shown here reflects the effect size (% signal change of the global mean) from the standard SPM mixed effects analysis corrected for multiple comparisons using an FDR of 0.05. Coronal slices of the cerebellum covering the extent of activation are shown in the left panel. The bottom panel shows activation in the basal ganglia.

Table 1

Summary of single pseudoword reading studies.

Study	Imaging modality	Production type
Petersen et al. (1990)	PET	Covert
Herbster et al. (1997)	PET	Overt
Rumsey et al. (1997)	PET	Overt
Brunswick et al. (1999)	PET	Overt
Fiez et al. (1999)	PET	Overt
Hagoort et al. (1999)	PET	Covert
Mechelli et al. (2000)	PET, fMRI	Covert
Riecker, Ackermann, Wildgruber, Meyer, et al. (2000)	fMRI	Overt
Tagamets et al. (2000)	PET	Covert
Lotze et al. (2000)	fMRI	Overt
Mechelli et al. (2003)	PET	Covert
Alario et al. (2006)	fMRI	Overt
Soros et al. (2006)	fMRI	Overt
Bohland and Guenther (2006)	fMRI	Overt

Table 2

Stimuli used in this experiment.

Stimulus type	Examples
Vowels	aa <i>as in bat</i> , ah (<i>hot</i>), ay (<i>hay</i>), ee (<i>beet</i>), eh (<i>bet</i>), ih (<i>bit</i>), o (<i>go</i>), oo (<i>boot</i>), uh (<i>but</i>)
CV syllables	baa, beh, bih, daa, dee, deh, dih, gaa, gah, geh, gih, guh, kaa, kah, kih, ko, kuh, pah, peh, po, puh, raa, rah, ree, reh, rih, ro, roo, ruh, taa, tay, tih, tuh
Bisyllables	bihko, pokeh, baateh, behbaa, behpah, behpeh, bihgih, daabeh, deebeh, deetah, dehdaa, dehdeh, dehpih, dihbuh, dihgah, dihpaa, dihtih, gaabih, gaakuh, gaapuh, gahpaa, gihteh, guhgih, kaadee, kahkaa, kaypah, kihbaa, kihpuh, kuhtih, pihdih, puhbih, puhgah, tahbaa, tahdee, tehpuh, tuhpa, tuhpeh

Region	Monosyllables > Baseline			Bisyllable > Monosyllable		
	x	y	z	x	y	z
	FDR = 0.05			FDR = 0.05		
Occipital_Inf_L				0.96 (3.1e-004)	-28	16
Occipital_Mid_R				2.04 (1.2e-004)	-16	-10
Occipital_Inf_R				1.03 (6.6e-004)	34	20
				1.50 (7.0e-005)	34	-16

Table 4
Results from the ROI analysis for the contrasts of interest.

Region	Monosyllables > Baseline		Bisyllable > Monosyllable	
	FDR = 0.05		FDR = 0.05	
	L	R	L	R
Frontal Lobe				
vPMC	1.20±0.54		0.43±0.28	
IFo	1.03±0.32		0.62±0.43	
IFt	0.51±0.19			
vMC	3.21±0.89	2.51±0.94	0.42±0.23	0.42±0.22
dMC	0.79±0.29	0.63±0.37		
preSMA	1.11±0.52	1.18±0.38	0.49±0.23	0.55±0.26
SMA	0.62±0.37	0.71±0.16		
FO	1.03±0.42	0.68±0.33		
aCO	1.34±0.71	1.04±0.54		
pCO	2.13±0.75	1.36±0.37		
PO	0.91±0.36	0.41±0.27		
aINS	0.94±0.35	0.52±0.21		
pINS	0.77±0.43	0.48±0.31		-0.20±0.13
aCG			-0.23±0.12	
pCG			-0.32±0.19	-0.32±0.17
aMFg		-0.45±0.19		
pMFg		-0.40±0.29		
FMC	-0.54±0.34			
FP	-0.34±0.22	-0.56±0.29	-0.35±0.16	-0.34±0.11
Temporal Lobe				
PP		0.90±0.64		
Hg	1.85±0.57	1.95±0.50	0.73±0.28	0.73±0.29
PT	1.31±0.64	1.42±0.43	0.83±0.30	0.59±0.24
aSTg		0.96±0.68	0.65±0.22	0.49±0.16
pSTg	0.79±0.61	1.13±0.53	0.79±0.32	0.53±0.28
pdSTs		0.66±0.32	0.38±0.26	
ITO	0.48±0.30		0.36±0.22	
adSTs		-0.45±0.31		
avSTs	-0.50±0.37	-0.67±0.39		-0.32±0.17
pvSTs	-0.36±0.25		-0.27±0.12	
aMTg		-0.56±0.30		
pMTg			-0.27±0.13	
MTO		-0.44±0.29		
TOF			0.46±0.27	
pTF		-0.30±0.12	0.32±0.21	
Parietal Lobe				
vSC	1.70±0.48	1.88±0.44		

Region	Monosyllables > Baseline		Bisyllable > Monosyllable	
	FDR = 0.05		FDR = 0.05	
	L	R	L	R
SPL			0.38±0.24	
pSMg	-0.46±0.25	-0.51±0.31		
Ag	-0.90±0.29	-1.01±0.21	-0.65±0.26	
PCN	-0.34±0.24	-0.49±0.23	-0.45±0.18	-0.26±0.15
Subcortical structures				
amCB	0.92±0.67	0.82±0.51		
splCB	0.32±0.25	0.48±0.30	0.27±0.17	
spmCB	0.76±0.39	0.85±0.38	0.44±0.21	0.55±0.27
Pal	0.91±0.14	0.62±0.18	0.31±0.14	0.28±0.17
Put	0.93±0.25	0.58±0.28	0.29±0.19	0.26±0.16
Tha	1.23±0.67	0.85±0.25		

Note. For each contrast, the effect sizes (% signal change of the global mean) of significantly active regions are shown together with their corresponding 95% confidence intervals (effect size ± confidence interval). Abbreviations: Ventral premotor cortex (vPMC), posterior dorsal premotor cortex (pdPMC), inferior frontal gyrus pars opercularis (IFo), inferior frontal gyrus pars triangularis (IFt), ventral motor cortex (vMC), dorsal motor cortex (dMC), presupplementary motor area (preSMA), supplementary motor area (SMA), frontal operculum (FO), anterior and posterior central operculum (aCO, pCO), parietal operculum (PO), anterior and posterior insula (aINS, pINS), anterior and posterior cingulate gyrus (aCG, pCG), anterior and posterior middle frontal gyrus (aMFg, pMFg), frontal medial cortex (FMC), frontal pole (FP), ventral and dorsal somatosensory cortex (vSC, dSC), superior parietal lobule (SPL), posterior supramarginal gyrus (pSMg), angular gyrus (Ag), precuneus (PCN), planum polare (PP), Heschl's gyrus (Hg), planum temporale (PT), anterior and posterior superior temporal gyrus (aSTg, pSTg), anterior and posterior dorsal superior temporal sulcus (adSTs, pdSTs), anterior and posterior ventral superior temporal sulcus (avSTs, pvSTs), inferior temporal occipital junction (ITO), anterior and posterior middle temporal gyrus (aMTg, pMTg), middle temporal occipital junction (MTO), temporal occipital fusiform gyrus (TOF), posterior temporal fusiform gyrus (pTF), anterior medial and lateral cerebellum (amCB, alCB), superior posterior medial and lateral cerebellum (spmCB, splCB), inferior posterior medial cerebellum (ipmCB), pallidum (Pal), putamen (Put) and thalamus (Tha).

Appendix

Lateralization test.

A nonparametric method based on bootstrapping (Efron, 1979) was used to test for lateralized effect sizes in a given ROI R for a given contrast comparing stimulus conditions A and B. The method consists of five steps; details are provided below.

Step 1: Calculate Distribution for Difference Between Conditions A and B in Each Hemisphere

For a given experimental run, data from condition A trials for each voxel in a given ROI (R) in the left hemisphere are represented in a matrix (M_A), with rows indexing the voxels in the ROI and columns indexing the trials corresponding to condition A in the given run. A vector ($vecA$) of the average activity for each voxel in the ROI is created by averaging across the columns of M_A . The same is done for condition B, providing a pair of vectors ($vecA$ and $vecB$) that represent the average activity for each voxel in ROI R in the left hemisphere in the two conditions. The difference between the activation in the two groups is then estimated by resampling from the voxel data vectors with replacement. A random voxel from $vecA$ and another random voxel from $vecB$ are selected, and the difference in their mean activations is recorded. This is repeated 5,000 times to generate a distribution with 5,000 samples that represents the difference in activity between condition A and condition B ($[A-B]_{Left}$) in the left hemisphere ROI. The same is done for the corresponding ROI in the right hemisphere giving $[A-B]_{Right}$.

Step 2: Calculate Distribution for Difference Between the Two Hemispheres

A second resampling is then performed to estimate hemispheric differences. A random sample from $[A-B]_{Left}$ (S_{Left}) and $[A-B]_{Right}$ (S_{Right}) is chosen, and the difference between the two samples is recorded (e.g., if we are testing left > right, we compute $S_{Left} - S_{Right}$). This is repeated 5,000 times to generate a distribution with 5,000 samples that represents the difference in activity between the two hemispheres for a single experimental run. This is done for each run, giving up to three distributions for a given subject.

Step 3: Calculate Distribution for Average Lateralization Across Runs for the Participant

A third resampling procedure is then performed to determine laterality effects for each participant. A random sample from each of the distributions representing the interhemispheric differences in each run for a given participant is chosen (S_1 , S_2 , up to S_3 depending on the number of runs), and the mean of the samples is recorded. The sampling procedure is repeated 5,000 times to generate a distribution with 5,000 samples that represents the average lateralization effect for a single participant. This is done for every participant.

Step 4: Calculate Distribution for Average Lateralization Across Subjects

A random selection from each participant's average lateralization effect distribution is then chosen, and the mean of these samples is recorded. This is repeated 5,000 times to generate a distribution with 5,000 samples that represents the average lateralization for the group.

Step 5. Calculate the Statistics for the Average Group Lateralization

Because specific hypotheses about the direction of lateralization were made, statistical testing for hemispheric differences consisted of computing the 90% confidence interval of the final distribution (average lateralization for the group). A 95% confidence interval would be used if one did not predict the direction of the lateralization a priori. The numerical limits of this confidence interval determine the statistical significance. If the limits span zero, then the null hypothesis cannot be rejected.
