

Research Article

Functional Parcellation of the Speech Production Cortex

Jason A. Tourville,^a Alfonso Nieto-Castañón,^a
Matthias Heyne,^a and Frank H. Guenther^{a,b,c}

Neuroimaging has revealed a core network of cortical regions that contribute to speech production, but the functional organization of this network remains poorly understood.

Purpose: We describe efforts to identify reliable boundaries around functionally homogenous regions within the cortical speech motor control network in order to improve the sensitivity of functional magnetic resonance imaging (fMRI) analyses of speech production and thus improve our understanding of the functional organization of speech production in the brain.

Method: We used a bottom-up, data-driven approach by pooling data from 12 previously conducted fMRI studies of speech production involving the production of monosyllabic and bisyllabic words and pseudowords that ranged from single vowels and consonant–vowel pairs to short sentences (163 scanning sessions, 136 unique participants, 39 different speech conditions). After preprocessing all data through the same pipeline and registering individual contrast maps to a common surface space, hierarchical clustering was applied to contrast maps randomly sampled from the pooled data set in order to identify consistent functional boundaries across subjects and tasks. Boundary completion was

achieved by applying adaptive smoothing and watershed segmentation to the thresholded population-level boundary map. Hierarchical clustering was applied to the mean within–functional region of interest (fROI) response to identify networks of fROIs that respond similarly during speech.

Results: We identified highly reliable functional boundaries across the cortical areas involved in speech production. Boundary completion resulted in 117 fROIs in the left hemisphere and 109 in the right hemisphere. Clustering of the mean within-fROI response revealed a core sensorimotor network flanked by a speech motor planning network. The majority of the left inferior frontal gyrus clustered with the visual word form area and brain regions (e.g., anterior insula, dorsal anterior cingulate) associated with detecting salient sensory inputs and choosing the appropriate action.

Conclusion: The fROIs provide insight into the organization of the speech production network and a valuable tool for studying speech production in the brain by improving within-group and between-groups comparisons of speech-related brain activity.

Supplemental Material: <https://doi.org/10.23641/asha.9402674>

Functional neuroimaging has revealed a core network of cortical areas in the brain that contribute to speech production (e.g., Basilakos, Smith, Fillmore, Fridriksson, & Fedorenko, 2017; Bohland & Guenther, 2006; Brown, Ingham, Ingham, Laird, & Fox, 2005; Eickhoff,

Heim, Zilles, & Amunts, 2009; Guenther, 2016; Indefrey, 2011; Simonyan, Ackermann, Chang, & Greenlee, 2016; Turkeltaub, Eden, Jones, & Zeffiro, 2002; Wise, Greene, Büchel, & Scott, 1999). Simple speech tasks, such as single-word reading, involve a number of cortical and sub-cortical brain regions, in addition to the primary motor cortex, including the medial and lateral premotor, somato-sensory, and auditory cortices; the anterior insula; and the cingulate motor area. Clinical and empirical evidence suggests these cortical areas fulfill distinct functional roles within this “speech production network,” and models of speech motor control themselves, derived from this evidence, hypothesize the existence of various integrated functional units contributing to speech production (Bohland, Bullock, & Guenther, 2010; Guenther, Ghosh, & Tourville, 2006; Hickok & Poeppel, 2007; Houde & Nagarajan, 2011).

^aDepartment of Speech, Language, and Hearing Sciences, Boston University, MA

^bDepartment of Biomedical Engineering, Boston University, MA

^cDepartment of Radiology, Massachusetts General Hospital, Charlestown

Correspondence to Jason A. Tourville: jtour@bu.edu

Editor-in-Chief: Ben Maassen

Received November 2, 2018

Revision received March 18, 2019

Accepted May 16, 2019

https://doi.org/10.1044/2019_JSLHR-S-CSMC7-18-0442

Publisher Note: This article is part of the Special Issue: Select Papers From the 7th International Conference on Speech Motor Control.

Disclosure: The authors have declared that no competing interests existed at the time of publication.

Our understanding of this functional organization, beyond broad classifications such as “motor” and “auditory” regions, however, remains limited, and it is unclear how the various regions of the speech network are organized into functionally homogenous “units.” Hindering our ability to address this question with functional neuroimaging may be a lack of power in prior studies due to small sample sizes and/or between-subjects anatomical and functional variability of the speech production network. Meta-analyses of multiple studies have been used to address small sample sizes, for example, activation likelihood estimation (Brown et al., 2005; Eickhoff et al., 2009; Turkeltaub et al., 2002). Such approaches provide valuable estimates of between-studies activation reliability but are limited in their ability to aggregate power across studies due to their reliance upon coordinate-based centers of group-level positive activations rather than subject- or group-level contrast maps (Costafreda, 2009). Furthermore, they do not allow for the analysis of co-activation patterns that can provide critical insights into functional subdivisions of cortical areas involved in speech production.

Cortical regions of interest (ROIs) have also been used to improve the statistical power of functional neuroimaging data (e.g., Nieto-Castañón, Ghosh, Tourville, & Guenther, 2003; Poldrack, 2007). ROIs based on observable macro-anatomical landmarks such as prominent cortical sulci (Caviness, Meyer, Makris, & Kennedy, 1996; Lancaster et al., 2000; Rademacher, Galaburda, Kennedy, Filipek, & Caviness, 1992; Talairach & Tournoux, 1988) are commonly applied and improve power by reducing between-subjects anatomical variability (Nieto-Castañón et al., 2003). The relationship between cortical anatomy and function, beyond the primary cortices, is highly variable, however, prompting the development of methods that define ROIs based on function rather than anatomy. Functional ROIs (fROIs) have been derived from assessing interregional resting-state functional connectivity (e.g., Cohen et al., 2008; Kim et al., 2010), but it remains unclear how relevant such task-free cortical parcellations are for studying task-specific processes in general and speech production in particular. Approaches based on task-based functional magnetic resonance imaging (fMRI) co-activation (e.g., Eickhoff et al., 2011) are more promising but, thus far, have relied on meta-analytic methods that suffer from relying on reported activity peak locations rather than full-brain activation patterns and on the aggregation of results across heterogeneous methods and analytic approaches.

In this article, we report on functionally homogenous ROIs in the speech production network that were derived by conducting an image-based mega-analysis (Costafreda, 2009) of data pooled from 12 previous fMRI studies conducted in our lab. More specifically, we reestimated surface-based speech–baseline contrasts for all participants and speech conditions using the same pipeline to ensure consistency across studies. We then used unconstrained between-subjects agglomerative hierarchical clustering to derive a population-level distribution of cortical boundary locations to determine the locations of functional boundaries that are reliable across a large sample of subjects and speech tasks (cf. Seghier & Price, 2009). Adaptive smoothing and watershed segmentation

were then applied to the population-level boundary distribution to form a set of fully bounded fROIs across the entire cortical surface in both cerebral hemispheres.

Finally, we investigated whether there were networks of fROIs that shared similar response patterns across our pool of subjects and speech tasks. Agglomerative hierarchical clustering was again applied, this time to the mean within-fROI speech–baseline responses. Clustering revealed nine distinct speech-positive cortical networks; these networks shed light on the organization of the cortical areas involved in speech production.

Method

Pooled Data Set

Data were pooled across 163 scanning sessions from 136 unique participants and 39 different speech conditions that were part of 12 previously conducted fMRI studies of speech production conducted over the past 15 years (see Table 1). All pooled data were acquired from neurologically normal, fluent, control participants. fMRI data were collected using gradient-echo MRI sequences. T1 structural scans (for registration to the FreeSurfer *fsaverage* space) were collected using MPAGE sequences. Blood oxygen level–dependent (BOLD) responses during speech were compared to those from a silent baseline task (viewing letter or symbol strings). Each contrast volume was then mapped to the standard cortical surface templates of both hemispheres separately, resulting in 581 pairs of cortical surface maps of speech-related activity.

Data Preprocessing

All data were reprocessed using the same analysis pipeline to ensure consistency across studies. Preprocessing was carried out using the CONN toolbox (Whitfield-Gabrieli & Nieto-Castañón, 2012) preprocessing modules (versions used: CONN17, SPM12). Each participant’s functional data were motion-corrected to their mean functional image and coregistered to their structural image. T1 volume segmentation and surface reconstructions were carried out using the FreeSurfer image analysis suite (freesurfer.net; Fischl, 2012; Fischl, Sereno, & Dale, 1999). Individual surfaces were then inflated to a sphere and coregistered to the FreeSurfer *fsaverage* template surface.

BOLD responses were high-pass filtered with a 128-s cutoff period and estimated at each voxel using a general linear model. The hemodynamic response function for each stimulus block was modeled using a canonical hemodynamic response function convolved with a boxcar function characterizing speech trials from each study. Model estimates for each speech condition and baseline were contrasted at each voxel to obtain speech–baseline contrast volumes.

Speech–baseline contrast volumes were resampled at the location of each corresponding subject’s cortical surface and projected to the two separate hemispheres of the *fsaverage* surface. The resulting surface contrast maps were then smoothed with 40 discrete diffusion steps (approximately

Table 1. Pooled functional magnetic resonance imaging data set: subject demographic and task summaries.

Study	Subjects	Speech stimuli (read aloud unless otherwise noted)	Baseline (silence)
Sequence learning ^a	12 (7F) Ages: 20–43	Monosyllabic pseudowords formed by legal or illegal consonant clusters	“xxx”
Sequence learning in persons who stutter ^b	13 (2F) Ages: 18–42	Monosyllabic pseudowords formed by legal or illegal consonant clusters	“xxx”
Syllable sequence representation ^c	18 (7F); Ages: 18–30 Fluent French Speakers	Bisyllabic pseudowords that varied by phonemic or suprasyllabic content	“XXXXX”
Syllable frame ^d representation	17 (9F) Ages: 20–43	Monosyllabic pseudowords that varied by phonemic, frame, or syllabic content	“xxx”
Consonant cluster ^e representation	16 (8F) Ages: 20–43	Bisyllabic pseudowords that varied by phonemic, cluster, or syllabic content	“xxx”
Overt production ^f	10 (3F) Ages: 19–47	Vowel (/V/), consonant–vowel (/CV/), or bisyllabic (/CVCV/) pseudowords	“xxxxx”
Auditory shift ^g	10 (6F) Ages: 23–36	Monosyllabic /CVC/ words under normal or altered auditory feedback (F1 shift)	“yyy”
Auditory category shift ^h	18 (9F) Ages: 19–33	Monosyllabic /CVC/ words under normal or altered auditory feedback (F1/F2 shift)	“***”
Somatosensory perturbation ⁱ	13 (6F) Ages: 23–51	/VV/ or /VCV/ pseudowords under normal or perturbed (interdental block) somatosensory feedback	“yyy”
Speech rate, clarity, and emphasis ^j	14 (7F) Ages: 18–35	Five-syllable sentences under fast, clear, emphatic, and normal conditions	Box characters
Covert vs. overt sequence production ^{k,l}	15 (8F) Ages: 21–33	Nonsense sequences of three syllables that varied by syllabic frame complexity	“*** **”
Sequence complexity ^m	13 (6F) Ages: 22–50	Sequences of three syllables that varied by sequence and syllabic complexity	“xxx xxx xxx”

Note. F = female.

^aSegawa et al. (2015). ^bBeal et al. (2012). ^cPeeva et al. (2010). ^dSegawa et al. (2013). ^eUnpublished. ^fGhosh et al. (2008). ^gTourville et al. (2008). ^hNiziolek & Guenther (2013). ⁱGolfopoulos et al. (2011). ^jGolfopoulos & Guenther (2011). ^kOverduin & Guenther (2009). ^lSpeech stimuli in this study were whispered. ^mBohland & Guenther (2006).

equivalent to a 10.8-mm full width at half maximum two-dimensional Gaussian smoothing kernel). This procedure resulted in 581 contrast maps characterizing BOLD responses across a wide range of subjects and different speech conditions.

Hierarchical Clustering

For every cortical vertex, we treated the 581 values of the pooled contrast maps at that location as a vector characterizing the heterogeneity of BOLD responses across different subjects and speech conditions. We then used a spatially constrained agglomerative hierarchical clustering algorithm to identify groups of neighboring vertices within the speech motor control network with similar BOLD response patterns (functionally homogeneous regions; cf. Seghier & Price, 2009). The procedure starts with the number of clusters equal to the number of cortical surface vertices and ends with a single cluster encompassing all vertices. At each step in the procedure, two clusters are joined among all possible pairs of adjacent clusters. The two clusters are chosen such that, when joined, the total resulting within-cluster variance in BOLD response patterns is minimized and the between-clusters variance is maximized (Ward, 1963). The map of boundaries between adjacent clusters is updated after each step. The procedure ends when only a single

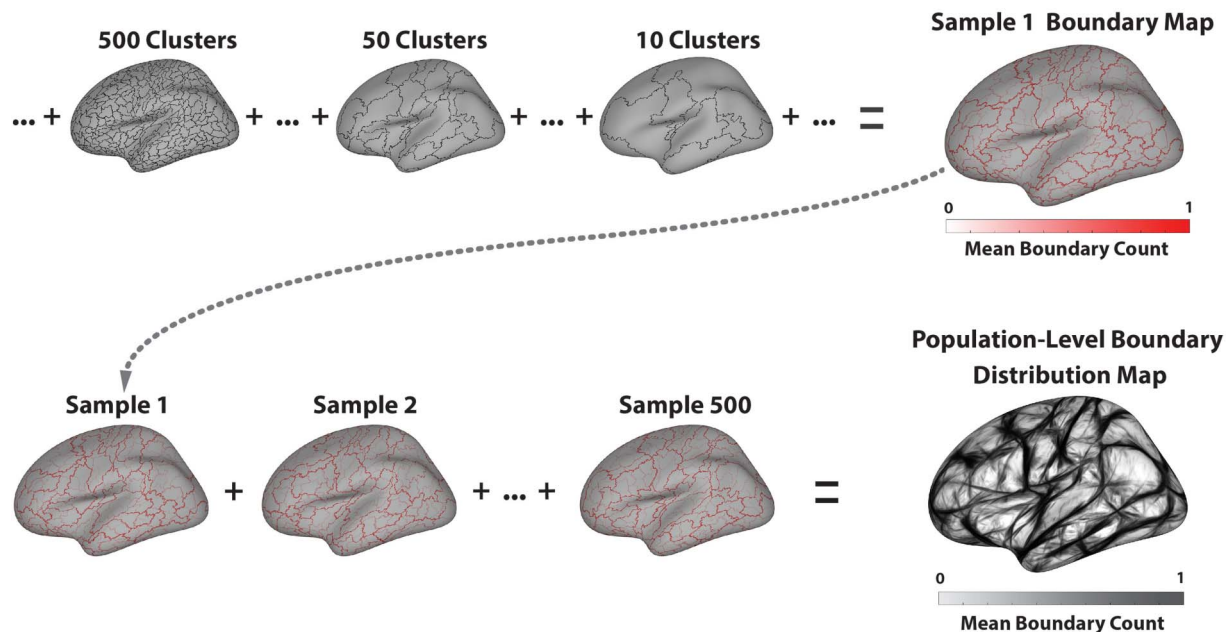
cluster (encompassing all vertices) remains. This is done within each hemisphere independently.

Finally, the boundary maps corresponding to the last 500 steps of the hierarchical clustering algorithm (parcellations ranging from 500 clusters to a single cluster) were averaged to produce a sample-level boundary distribution map, where higher values indicate boundaries that persisted longer in the clustering process and therefore divide regions within the speech motor control network that are more functionally distinct (see Figure 1, top).

Boundary Reliability

The hierarchical clustering procedure above was repeated 500 times, each performed on a random resampling of the data set, in order to characterize the variability of the resulting boundary locations. For each repetition, we randomly selected, with replacement, a set of 581 pooled contrast maps from the entire set of 581 pooled contrast maps (Efron, 1982). This form of resampling was chosen to obtain robust estimates of population-level parameters (the location of functional boundaries in this case) from a limited-sample subset. For each of these 500 randomly selected samples of 581 contrast maps, we performed the same spatially constrained agglomerative hierarchical

Figure 1. (Top) Illustration of the process for deriving a sample boundary distribution map. Only the left hemisphere is shown for simplicity. The boundary map from each step of the hierarchical clustering process is summed and divided by the total number of clustering steps (the number of initial vertices), resulting in a mean sample-level boundary map (upper right plot). Higher values in this map (darker lines) denote vertices that were marked as boundaries earlier in the clustering process; that is, they divide more functionally distinct regions than lighter lines. (Bottom) Illustration of the process for building the population-level distribution. The sample boundary maps are summed and divided by the total number of samples (500), resulting in the population-level boundary distribution map (lower right plot). Higher values in this map (darker lines) denote vertices that more reliably divide functionally distinct regions across the samples.



clustering procedure to obtain a new sample-level boundary distribution. The resulting 500-sample-level boundary distribution maps were averaged to build a population-level boundary distribution map (see Figure 1, bottom). Finally, this map was spatially processed using an adaptive smoothing kernel in order to increase the image contrast along locally linear boundaries.

Boundary Statistics

The entire procedure, including 500 random resampling repetitions, was finally repeated once again but now starting with a null data set of 581 random pooled contrast maps. This null data set was generated by independently assigning to each vertex, for each subject and speech condition, a random value from a Gaussian distribution with zero mean and unit variance. The resulting null data set pooled contrast maps were then spatially smoothed using 64 discrete diffusion steps in order to approximate the level of spatial covariance between adjacent vertices observed in our real data set.

For each vertex, its population-level boundary distribution value (obtained from the real data set) was ranked against the distribution of population-level boundary values obtained from the null data set in order to compute an uncorrected p value (defined as the percentage of null data set values equal or above the observed real data set value at each vertex). Finally, these p values were corrected

for multiple comparisons across the entire cortical surface using false discovery rate (FDR; Benjamini & Hochberg, 1995) to construct a map of FDR-corrected p values characterizing the reliability of the resulting boundary locations.

Boundary Completion

Watershed segmentation was applied to the population-level boundary map to form a set of fully bounded fROIs. Surface-level iterative smoothing (Hagler, Saygin, & Sereno, 2006) was applied to the population-level boundary map prior to segmentation in order to remove spurious boundaries. To determine the appropriate level of smoothing (number of diffusion steps), we applied the segmentation on a range of smoothing levels and chose the resulting fROI parcellation that featured boundaries that best matched the original population-level boundary map thresholded at $p_{\text{FDR}} < .001$. Smoothing the original map using 40 discrete diffusion steps resulted in the best match. Last, all fROI masks resulting from the watershed procedure were further cleaned using a sequence of five binary erosion steps followed by binary dilation steps in order to remove small or thin parcels (fROIs with radius smaller than five vertices, approximately 4 mm).

Region Labeling

To provide a convenient means of referencing specific fROIs, each resulting fROI was given a label according to

its overlap with a prominent macro-anatomical landmark. To differentiate fROIs that overlapped the same landmark, a number suffix was added to each ROI. For instance, an fROI lying primarily on the cingulate gyrus was labeled *cg.1*, and another region also primarily on the cingulate gyrus was labeled *cg.2*. However, a nearby region that spans the cingulate sulcus, thereby encompassing portions of cingulate and superior frontal gyrus, was labeled *cgs.1*. Labels in the left and right hemispheres were assigned independently; that is, *cg.1* in the left hemisphere is not necessarily the contralateral homologue of *cg.1* in the right hemisphere.

Identifying Functional Networks

To investigate the degree of speech response similarity between fROIs, we computed the average speech–baseline responses aggregated across all vertices within each fROI. When computing these average response patterns, we explicitly disregarded surface vertices that laid within the FreeSurfer *fsaverage* cerebral medial mask (representing points in each hemisphere surface reconstruction that do not correspond with cortical areas), as well as fROIs where more than 70% of their vertices laid within the same medial mask. We then applied unconstrained agglomerative hierarchical clustering (Ward’s minimum variance method), this time across fROIs rather than across voxels/vertices as we did originally, in order to cluster the fROIs according to the similarities and differences of their mean speech response patterns. The unthresholded dendrogram resulting from this post hoc clustering was used to identify networks (groups of speech fROIs, not necessarily contiguous) that shared similar response patterns across the different speech conditions in our database. Networks were identified based on the level in the dendrogram at which a cluster of fROIs formed its own branch and with consideration of anatomical or presumed functional contributions of the fROIs within a larger cluster. To enable visualization of these networks, each fROI was assigned a color such that proximal regions in the dendrogram were assigned similar colors.

Finally, to characterize the functional relationship between speech fROI networks, we calculated the “response distance” between all pairs of networks. We first calculated the Euclidean distance between the mean speech–baseline response patterns from all 581 speech contrasts for all pairs of individual fROIs. The response distance between two networks was then given by the average distance between all fROI pairs in each of those two networks.

Results

The location of cortical vertices that are significantly likely ($p_{\text{FDR}} < .001$) to demark a population-level functional transition during speech production is shown in red in the top panel of Figure 2, along with the completed speech fROI boundaries (thick black lines in both panels of Figure 2). The cortical surface of the left hemisphere was divided into 117 fROIs, while the right hemisphere was parcellated into 109. Some fROIs along the hemispheric margins are not

visible in Figure 2; a set of surface maps with additional views showing the complete labeled fROI parcellation is provided in Supplemental Material S1.

fROI Boundaries: Relation to Cortical Topography

Significant boundaries were found primarily along the crowns of prominent gyri that are active during speech production. Notably, the fundus of the central sulcus, the putative division of primary motor and primary somatosensory cortex (Brodmann areas 4 and 3, respectively), was not found to represent a significant functional boundary (see Figure 2, bottom). Nor was there a significant functional distinction along the fundi of the precentral, postcentral, and superior temporal and cingulate sulci, anatomical landmarks that regularly mark regional boundaries in cortical parcellations (e.g., Caviness et al., 1996; Desikan et al., 2006; Tourville & Guenther, 2012; Tzourio-Mazoyer et al., 2002). Sulcal fundi that did form reliable functional boundaries within the speech network include the posterior Sylvian fissure, dividing putative primary and secondary auditory regions (Heschl’s gyrus and planum temporale) from the somatosensory opercular cortex. Also within the Sylvian fissure, a significant boundary ran along the first transverse sulcus, marking a division of primary auditory cortices from more anterior higher order auditory cortices of the planum polare.

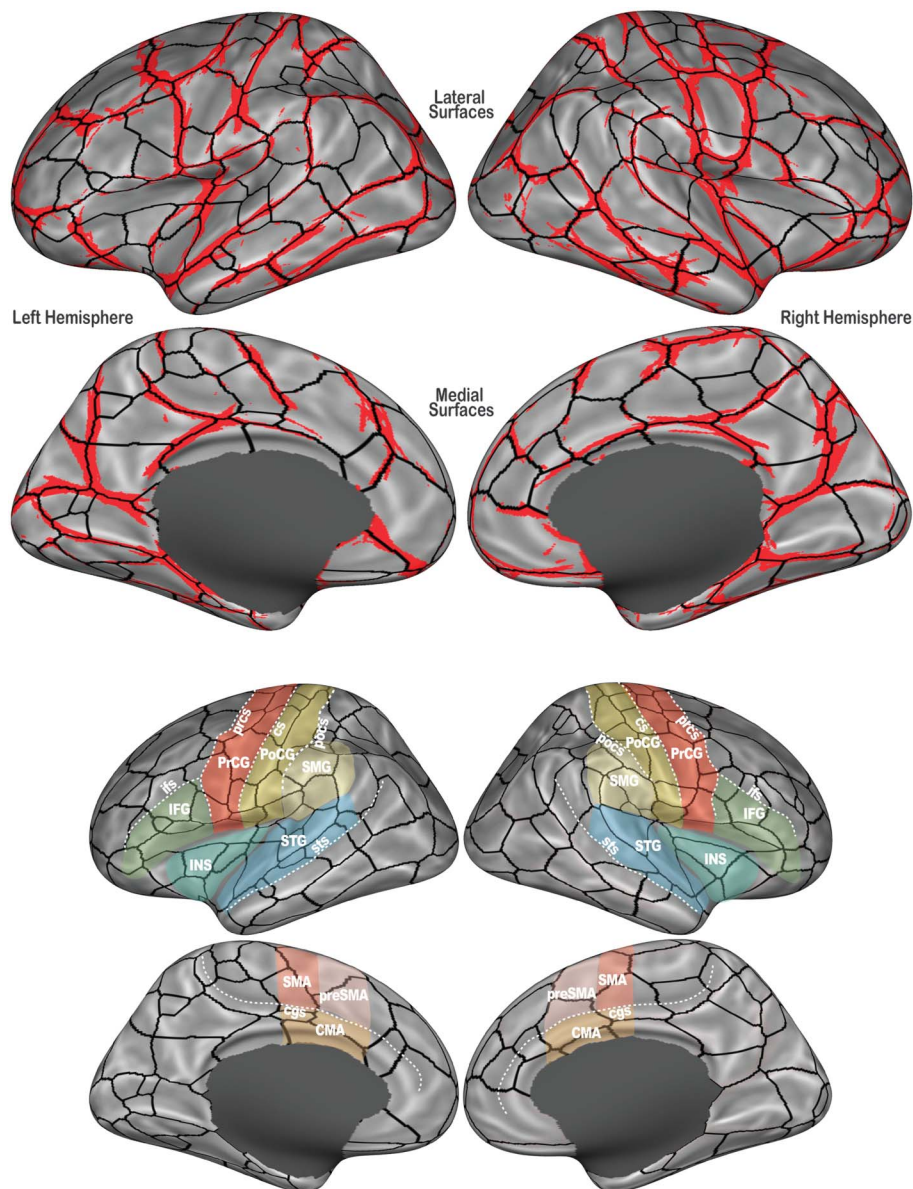
Exceptions to the gyrus-oriented divisions within the speech network include a split of the precentral sulcus into dorsal and ventral segments bilaterally. This functional distinction lies lateral to the “knob” in the caudal bank of the precentral gyrus (e.g., Yousry et al., 1997) that marks the transition from the hand motor area more medially to the vocal articulator representations more laterally. Significant boundaries also subdivided the medial prefrontal wall, forming borders that lie near the presumed anatomical bounds of the combined supplementary and presupplementary motor areas (SMA and preSMA, respectively). In the right hemisphere only, a significant border bisects the ventral precentral sulcus.

The boundary completion step formed fully bounded fROIs (black lines in Figure 2) according to the reliability of boundary locations, that is, overlap in the population-level boundary map. In addition to borders aligned with the highly significant boundary locations, roughly orthogonal borders that follow subthreshold “peaks” in the population-level boundary map subdivide the prominent speech-positive gyri. For instance, the lateral sensorimotor cortex is segmented into several fROIs along its dorsal–ventral extent. Likewise, the superior temporal gyrus is divided into several anterior–posterior fROIs.

fROI Boundaries: Relation to BOLD Response

Figure 3 shows the fROIs overlaid on areas that are significantly more active during speech than during baseline tasks in our sample. The pooled speech–baseline contrast t map was initially thresholded at a vertex

Figure 2. (Top) The location of significantly likely functional boundaries is shown in red (population-level boundary map thresholded at $p_{FDR} < .001$) and completed functional region of interest (fROI) boundaries following watershed segmentation are shown in black on the inflated FreeSurfer *fsaverage* cortical surface template. The lateral (top) and medial (bottom) surfaces of the left (left) and right (right) hemispheres are visible. The grayscale surface shading indicates cortical topography: Bright shading indicates the convex curvature of gyral crowns; darker shading indicates the concave curvature of sulcal depths. The dark-filled gray area on each medial surface masks the noncortical region of the cerebral medial wall. (Bottom) The completed fROI boundaries are shown again with prominent sulci (dotted lines) and cortical regions (color-filled regions) involved in speech production labeled. Sulcus abbreviations: cgs = cingulate sulcus; cs = central sulcus; ifs = inferior frontal sulcus; pocs = postcentral sulcus; prcs = precentral sulcus; sts = superior temporal sulcus. Cortical region abbreviations: CMA = cingulate motor area; IFG = inferior frontal gyrus; INS = insula; PoCG = postcentral gyrus; PrCG = precentral gyrus; preSMA = presupplementary motor area; SMA = supplementary motor area; SMG = supramarginal gyrus; STG = superior temporal gyrus.

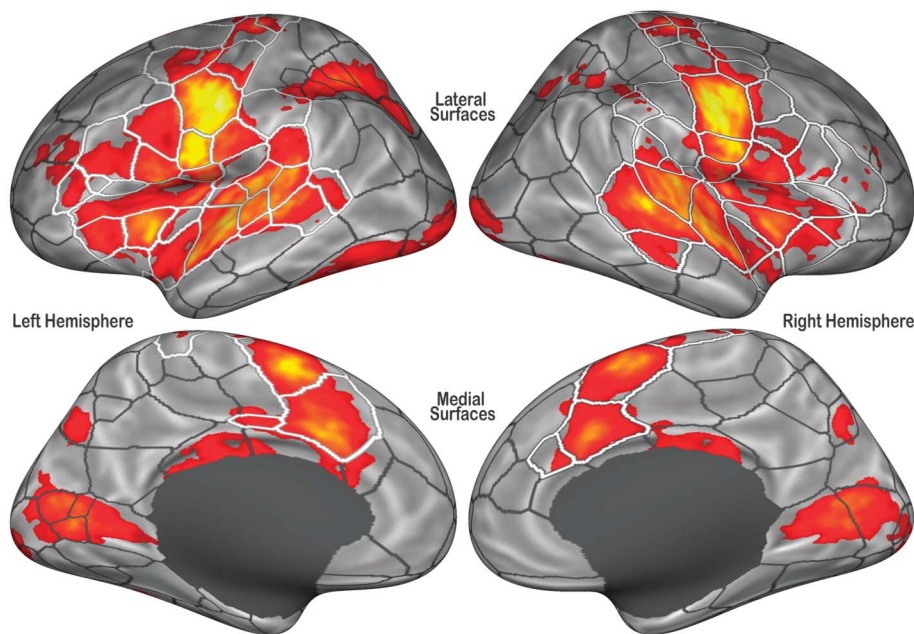


significance level of $p < .001$. Monte Carlo simulations were then run to estimate cluster-level significance thresholds (Hayasaka & Nichols, 2003), and the t map was thresholded to ensure a cluster-wise FDR (p_{FDR}) $< .05$.

To better illustrate the relationship between speech-related BOLD responses and the fROIs, the boundaries are

overlaid on the unthreshold pooled speech-baseline t map in Figure 4. The sign and level of responses are generally consistent within each fROI: Large changes in activation (average speech response across all tested conditions) are seen across fROI borders but not within them. Large contiguous clusters of activation (e.g., along the superior temporal

Figure 3. Completed functional region of interest (fROI) boundaries overlaid upon a t map of cortical vertices that were significantly more active during speech production compared to baseline (vertex-level threshold: $p < .001$; cluster-level correction: $p_{FDR} < .05$). The boundaries of fROIs in cortical areas that are reliably active during speech production are highlighted in white.



gyrus and lateral central sulcus) are subdivided into several smaller fROIs with the activation within each division generally uniform. This is generally the case even for clusters of negative activation that lie outside the speech network. A similar relationship between the fROI boundaries and speech-related changes in brain activation is also seen in areas of the cortex in which activity is reduced during speech production. For instance, a large bilateral cluster of negative activation in the vicinity of the angular gyrus is subdivided into three subregions in both hemispheres.

There are exceptions to the uniformity of the speech–baseline responses within the completed fROIs. In both hemispheres, a large fROI encompasses much of the frontal and central operculum and adjacent dorsal insula. In the left hemisphere, this region spans multiple activation peaks from apparently independent response clusters (arrowheads labeled *f* and *g* in Figures 4A and 4C). Another exception can be seen bilaterally along the dorsal central sulcus where a small cluster of positive activity, hypothesized to be associated with respiratory control during voicing (e.g., Guenther et al., 2006; Takai, Brown, & Liotti, 2010), is present. In both hemispheres, the cluster is divided into small fROIs, some of which include a mix of strongly positive and weakly negative speech-related activity (see Figure 4A, arrowheads *h* and *i*).

In Figure 4B, the speech–baseline contrast map is shown relative to a parcellation of the cortex that is based on cortical anatomy, rather than function. These structural ROI (sROI) boundaries are based on macro-anatomical landmarks, including prominent sulci and gyri that are commonly used in brain atlases (e.g., Duvernoy, 1999;

Petrides, 2014) to label cortical regions (e.g., precentral and central sulcus bounding the precentral gyrus, the superior temporal sulcus and Sylvian fissure bounding the superior temporal gyrus). The sROI boundaries shown in Figure 4B are a modification of the parcellation system developed by Caviness et al. (1996) and later adapted as a cortical ROI atlas for use with FreeSurfer (Desikan et al., 2006). The modifications were designed to capture key functional divisions in the cortical speech motor network based on neuroimaging and physiological mapping studies (Tourville & Guenther, 2012; see Peeva et al., 2010; Cai et al., 2014, for examples of the application of this cortical labeling system in the analysis of speech neuroimaging data). So, while not explicitly based on functional divisions in the cortex, the sROIs approximate those divisions, resulting in the relatively high response uniformity within the sROIs evident in Figure 4B.

A closer inspection of within-ROI responses reveals a greater mix of response level and direction in the sROIs in some key speech regions compared to the fROIs. For instance, the arrowhead labeled *b* in Figure 4 marks the peak of a distinct cluster of activity in left ventral precentral and adjacent posterior inferior frontal cortex, an area hypothesized to encode speech motor commands (Tourville & Guenther, 2011). The functional parcellation isolates the majority of this cluster within a single fROI (*l.prcs.I*) and isolates it from the stronger activity in lateral orofacial somatomotor cortex (*l.cs.I*) and other adjacent positive (cluster labeled *a* in Figure 4C, top) and negative (cluster labeled *c*) response clusters. The structural parcellation, on the other hand, distributes the same cluster across three

different sROIs (ventral premotor cortex, dorsal inferior frontal gyrus, and ventral inferior frontal gyrus; see Figure 4C, bottom). The ventral premotor cortex extends dorsally to include a portion of cluster *a*, while dorsal inferior frontal gyrus and ventral inferior frontal gyrus extend rostrally to include the cluster of weakly negative responses (labeled *c*).

Two other key left-hemisphere speech regions with responses that are more uniformly distributed within the fROIs than the sROIs are highlighted by the arrowheads labeled *d* and *e* in Figure 4. The structural parcellation mixes negative and positive activity (arrowhead *d*) in the parietal operculum (see Figure 4C, bottom); the functional parcellation isolates this cluster of negative activity (fROI *l.op.3* in Figure 4C, top) from surrounding positive activity, resulting in more uniform responses in adjacent opercular and superior temporal fROIs, an area thought to represent the interface between the auditory and motor systems involved in speech. On the medial surface, a cluster of activation (arrowhead *e*) centered at the border of the SMA and preSMA sROIs (see Figure 4D, bottom) is isolated within a single fROI (*l.sfg.5*; see Figure 4D, top). This area is hypothesized to contribute to the initiation and timing of speech motor commands.

Functional Organization of the Speech Production Network

Hierarchical clustering of the fROIs based on their mean within-ROI response from all 581 speech–baseline contrasts revealed a core set of fROIs spanning the lateral central sulcus and superior temporal gyrus bilaterally, that is, orofacial somatomotor and auditory cortex (dark red ROIs in Figure 5; region labels highlighted in red in the dendrogram in Supplemental Material S2) that responded similarly across a wide range of speakers and speech conditions. This core sensorimotor network formed its own branch in the clustering dendrograms. fROIs outside the core sensorimotor network clustered into two broad categories: those that also exhibited positive mean responses during overt speech production compared to baseline (“speech-positive” fROIs; regions filled in shades of green, yellow, or orange in Figure 5) and those with a negative mean response (“speech-negative” fROIs; regions filled in shades blue in Figure 5). The fROIs in each of the speech-positive networks are listed in Table 2.

Excluded from the core sensorimotor network are fROIs on the more dorsal portion of precentral gyrus that divide activation that is commonly attributed to breathing control (e.g., Guenther et al., 2006; Takai et al., 2010). These fROIs were clustered together in a bilateral dorsal somatomotor network.

Surrounding much of the core sensorimotor network is a bilateral cluster of fROIs that has the shortest response distance from the core network (see Table 3). Within this primary flanking network, fROIs in the bilateral posterior superior temporal sulcus and the posterior portion of right planum temporale form a distinct cluster of higher order posterior auditory processing regions (see Supplemental

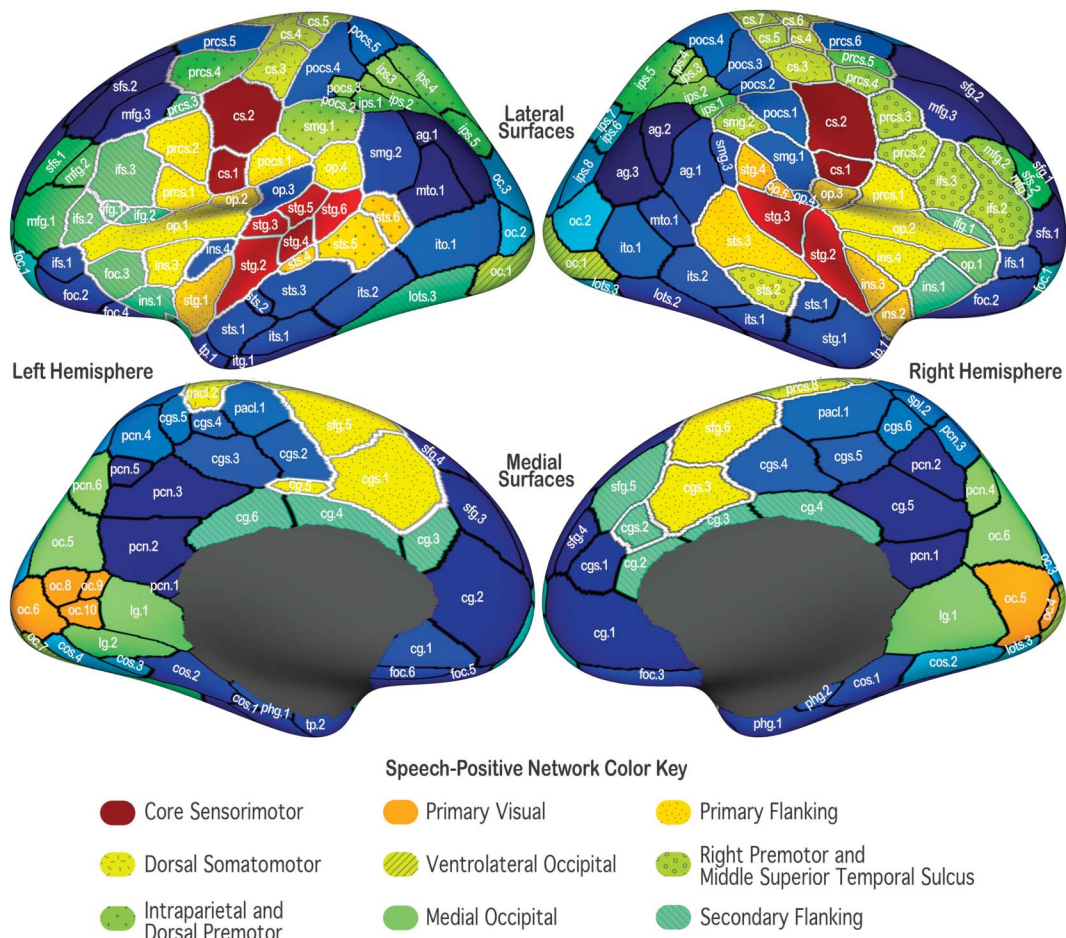
Material S2). fROIs in the bilateral medial anterior superior temporal gyrus and adjacent insula form another distinct subcluster from the remaining ROIs in the network. Of these, a small set of fROIs in the bilateral posterior central operculum and left ventral postcentral gyrus formed a subcluster of somatosensory processing regions. The remaining fROIs in the network formed a larger cluster that includes the bilateral ventral premotor cortex and anterior insula, frontal and anterior central operculum, SMA and preSMA, cingulate motor area, and posterior-most portion of left parietal operculum and planum temporale. Within this cluster, a left-lateralized subcluster of fROIs in the premotor cortex (*l.prcs.1* and *l.prcs.2*) and anterior insula (*l.ins.3*) is present.

Moving further outward from the core sensorimotor network along the cortical surface, a secondary flanking network comprising fROIs lying anterior and medial to much of the primary flanking network is present. This network includes fROIs bilaterally in the inferior frontal gyrus, including the orbital portion of the gyrus and adjacent fROIs in the frontal operculum and the anterior insula, and the middle and anterior segments of the cingulate gyrus. Whereas the majority of the left inferior frontal gyrus is included in this network, only a small portion of the ventral portion of the gyrus in the right hemisphere is included. A small fROI near the junction of the inferior frontal and precentral sulci in only the left hemisphere is also included in this network. Distant from the other regions in this network, a large fROI spanning the posterior portion of the lateral occipitotemporal sulcus in the left hemisphere (*l.lots.3*) was also included in the secondary flanking network.

The secondary flanking network has a unique functional relationship with the other networks formed by the hierarchical clustering, being the shortest response distance from most of six of the other eight speech-positive networks that were formed (see Table 3). This includes the primary flanking and dorsal somatomotor networks mentioned above, a bilateral intraparietal sulcus and dorsal premotor cortex network, and a unilateral network of fROIs along the right premotor and posterior inferior frontal cortex, middle superior temporal sulcus, and supramarginal gyrus. It is also the shortest response distance from two different bilateral clusters of higher order visual processing regions: the ventrolateral occipital network along the posterior portion of the inferior occipital gyrus and the medial occipital network, which includes fROIs in the lingual gyrus and cuneal cortex extending into the parieto-occipital sulcus (see Table 3). The latter has the shortest response distance from the secondary flanking network and was closest to the secondary flanking network in the hierarchical clustering dendrogram (see Supplemental Material S2).

Forming its own branch among the speech-positive fROIs outside the core sensorimotor network are fROIs that subdivide activation found within the posterior portion of the calcarine sulcus bilaterally (the V1 network). The primary visual cortex is consistently mapped to this area (e.g., Hinds et al., 2008; Van Essen, Drury, Joshi, & Miller, 1998), implying greater low-level visual processing of stimuli during the speech task than baseline.

Figure 5. Functional networks of speech functional regions of interest (fROIs). fROIs were grouped using hierarchical clustering of the average within-ROI response for all speech–baseline contrasts. To illustrate fROIs with similar response patterns, colors were assigned to each fROI according to its position in the clustering dendrogram (see Supplemental Material S2). fROIs filled with similar/dissimilar colors exhibit similar/dissimilar speech response patterns. Patterned stippling was overlaid on some networks to improve differentiation. fROIs that are consistently active during speech production are highlighted in white outlines (cf. Figure 3).



Prominent clusters were also formed by regions that showed reduced BOLD response during speech compared to baseline. Notably, areas typically associated with the default-mode resting-state network, including angular gyrus, dorsolateral and anterior medial prefrontal cortex, inferior frontal cortex, and precuneal cortex, clustered together (cf. Buckner, Andrews-Hanna, & Schacter, 2008; Lee, Smyser, & Shimony, 2013).

Discussion

In this study, we combined hierarchical agglomerative clustering and watershed segmentation to parcellate the cerebral cortex into fROIs that respond similarly during speech production. This effort to identify functionally homogenous regions was undertaken with two goals in mind: to improve group-level comparison of functional imaging data and to improve our understanding of the functional organization of speech-responsive cortices.

To meet these goals, the derived region boundaries must be robust across speakers and speaking conditions. To ensure this, the clustering was performed on speech BOLD responses from 136 unique speakers producing a wide range of speech tasks that included overt and covert production, normal and perturbed auditory feedback, and both native words and pseudowords that varied in terms of phonological, phonetic, and articulatory complexity and familiarity.

The fROIs derived here provide a complete parcellation of the cortex that minimizes speech BOLD response variability and therefore has the potential to improve the detection of speech-related BOLD effects by providing a better means of aligning functionally uniform regions across subjects. The parcellation system is available as labeled FreeSurfer *fsaverage* surfaces (available upon request from the authors); it can be easily applied to individual data sets that have been processed through the FreeSurfer surface reconstruction pipeline.

Table 2. A list of the functional regions of interest (fROIs) in each speech-positive network identified by hierarchical clustering of the mean fROI speech response pattern.

Network	Left-hemisphere fROIs	Right-hemisphere fROIs
Core sensorimotor	l.cs.1, l.cs.2, l.stg.2, l.stg.3, l.stg.4, l.stg.5, l.stg.6	r.cs.1, r.cs.2, r.stg.2, r.stg.3
Primary visual	l.oc.6, l.oc.8, l.oc.9, l.oc.10	r.oc.4, r.oc.5
Primary flanking	l.cg.5, l.cgs.1, l.ins.3, l.op.1, l.op.2, l.op.4, l.pocs.1, l.prcs.1, l.prcs.2, l.sfg.5, l.stg.1, l.sts.4, l.sts.5, l.sts.6	r.cgs.3, r.ins.2, r.ins.3, r.ins.4, r.op.2, r.op.3, r.op.5, r.prcs.1, r.sfg.6, r.stg.4
Dorsal somatomotor	l.cs.3, l.cs.4, l.cs.5, l.cs.6, l.pacl.2, l.prcs.6	r.cs.3, r.cs.4, r.cs.5, r.cs.6, r.cs.7, r.prcs.8
Ventrolateral occipital	l.oc.1, l.oc.7	r.oc.1
Intraparietal and dorsal premotor	l.ips.1, l.ips.2, l.ips.3, l.ips.4, l.ips.5, l.pocs.2, l.pocs.3, l.prcs.4, l.smg.1	r.ips.1, r.ips.2, r.ips.3, r.ips.4, r.ips.5, r.prcs.5
Right premotor and middle superior temporal sulcus		r.ifs.2, r.ifs.3, r.mfg.1, r.mfg.2, r.prcs.2, r.prcs.3, r.prcs.4, r.smg.2, r.sts.2
Medial occipital	l.lg.1, l.lg.2, l.oc.5, l.pcn.6	r.lg.1, r.oc.6, r.pcn.4
Secondary flanking	l.cg.3, l.cg.4, l.cg.6, l.foc.3, l.ifg.1, l.ifg.2, l.ifs.2, l.ifs.3, l.ins.1, l.ins.2, l.lots.3, l.mfg.1, l.mfg.2, l.prcs.3, l.sfs.1	r.cg.2, r.cg.3, r.cg.4, r.cgs.2, r.ifg.1, r.ins.1, r.op.1, r.sfg.5, r.sfs.2

Note. cg = cingulate gyrus; cgs = cingulate sulcus; cos = collateral sulcus; cs = central sulcus; foc = frontal orbital cortex; fp = frontal pole; ifg = inferior frontal gyrus; ifs = inferior frontal sulcus; ins = insula; ips = intraparietal sulcus; itg = inferior temporal gyrus; ito = inferior temporal-occipital area; its = inferior temporal sulcus; lg = lingual gyrus; lots = lateral occipitotemporal sulcus; mfg = medial frontal gyrus; oc = occipital cortex; op = pars opercularis; pacl = paracentral lobule; pcn = precuneus; phg = parahippocampal gyrus; pocs = postcentral sulcus; prcs = precentral sulcus; sfg = superior frontal gyrus; sfs = superior frontal sulcus; smg = supramarginal gyrus; spl = superior parietal lobule; stg = superior temporal gyrus; sts = superior temporal sulcus; tp = temporal pole.

Table 3. Mean response distances between pairs of speech-positive networks identified by post hoc hierarchical clustering of mean functional region of interest (fROI) responses.



	Core	V1	Flank1	dS-M	vIOC	RPMC-midSTS	IPS-dPMC	mOC	Flank2
Core	1317	3227	2621	3378	3734	3520	3630	3536	3650
V1		573	1330	1646	1236	1701	1488	911	1405
Flank1			693	907	1154	942	885	833	803
dS-M				540	1162	860	786	801	751
vIOC					303	1078	1056	956	952
RPMC-midSTS						594	739	876	706
IPS-dPMC							477	762	647
mOC								319	641
Flank2									496

Note. In each column, the row highlighted in yellow indicates the network with the shortest response distance from the network listed in the column heading; the row highlighted in red indicates the network with the greatest response distance from the network listed in the column heading. Values along the diagonal highlighted in gray represent the mean response distance between all fROIs within the network listed in the column heading. Above the table, a simplified version of the clustering dendrogram illustrates the nesting of speech-positive networks (branch height is arbitrary; branches within each network and those to speech-negative fROIs were eliminated; see Supplemental Material S2 for complete dendrogram). Colors in the headings indicate the approximate midpoint of the range of color assigned to regions in that network in Figure 5 and Supplemental Material S2. The asterisk (*) indicates the omitted branch to the default mode and other speech-negative fROIs. Core = core sensorimotor; Flank1 = primary flanking; V1 = primary visual cortex; dS-M = dorsal somatomotor; RPMC-midSTS = right premotor and middle superior temporal sulcus; mOC = medial occipital cortex; Flank2 = secondary flanking; IPS-dPMC = intraparietal and dorsal premotor cortex; vIOC = ventrolateral occipital cortex.

Given that our goal was to maximize the homogeneity of speech responses within the parcels derived from clustering, we used Ward's minimum variance clustering algorithm. A suitable quantitative measure of regional homogeneity is the variance of speech responses within each parcel. Ward's algorithm optimizes this measure; specifically, it minimizes within-cluster variance while maximizing between-clusters variance. Alternative forms of clustering, for example, single- or complete-linkage clustering methods, do not necessarily optimize the same measures. Ward's algorithm also offers better performance than other clustering algorithms when applied to simulated and real fMRI data, representing a good compromise between the spatial consistency provided by spectral clustering and the realistic representation of plausible functional patches produced by k-means clustering (Thirion, Varoquaux, Dohmatob, & Poline, 2014; cf. Park et al., 2017). However, a challenge for Ward's and other clustering approaches is how to best determine the appropriate number of parcels to be included in the parcellation system. How many regions is the right number? We avoided this challenge by iteratively applying unconstrained clustering (i.e., clustering continued until only one region, encompassing the entire cortex, remained) to random samples of a large set of overt speech production fMRI data. The boundaries from each sample were aggregated to build a population-level map of the likelihood of a boundary at every vertex of the cortical surface. An image segmentation algorithm was then applied to this map to form fully enclosed parcel boundaries that were constrained to lie along local maxima of the boundary likelihood. Thus, the number of fROIs and their borders are governed by the boundary likelihoods rather than a predetermined threshold.

Perhaps more importantly, the population-level likelihood map allowed us to identify the location of significant functional boundaries in the cortex during speech production. The areas in red at the top of Figure 2 indicate the location of transitions in the pattern of BOLD responses that are highly consistent across a wide range of speech production tasks and speakers. Within the core speech production cortices (areas highlighted at the bottom panel of Figure 2), these transitions are largely symmetric across the two hemispheres. Functional boundaries were found along the precentral, postcentral, and superior temporal gyri; notably, these significant boundaries were found along gyral crowns rather than in the depths of the sulci that mark the borders of these gyri.

Functional boundaries along gyral crowns are not a surprise, of course. Cytoarchitectonic transitions along the crowns of the precentral and postcentral gyri, for instance, and along the hemispheric margin of the superior temporal gyrus have been observed by Brodmann and many others (see Zilles & Amunts, 2010). The alignment of these microstructural and functional divisions prompted us to incorporate approximations of those boundaries in the structural labeling system that we developed for analysis of speech neuroimaging studies shown in Figure 4B (Tourville & Guenther, 2012).

What is surprising, however, is the paucity of functional transitions along prominent sulci. This lack of functional

boundaries along sulcal fundi could reflect biases in the BOLD signal rather than true functional-anatomical relationships. For instance, large draining veins can effectively reduce the spatial resolution of the BOLD signal from within sulci (e.g., Wilson, 2014), an effect that may be exacerbated by partial voluming across adjacent sulcal banks. While our choice of a surface-based preprocessing pipeline is explicitly designed to minimize the chance of unintended BOLD signal mixture across adjacent sulcal banks (regions that may be close in the three-dimensional volume but relatively distant along the cortical surface), such spatial biases could contribute to the homogeneity of responses across sulcal banks observed here.

Additional tasks and scans and alternative acquisition methods can mitigate this spatial bias (Wilson, 2014). Such countermeasures may not be practical for all studies, however, and are incomplete. Additional intrinsic biases in the BOLD signal related to the cerebral vasculature and that vary as a function of the orientation and depth of the cortex (Gagnon et al., 2015; R. S. Menon, 2012; Viessmann, Bianciardi, Scheffler, Wald, & Polimeni, 2018) are a subject of ongoing research. Though the functional boundaries observed here may be spatially biased, they do reflect speech-related responses derived from common fMRI acquisition and data processing methods. As such, they are a more appropriate means of analyzing and/or interpreting speech fMRI data than an unbiased "ground truth" functional parcellation of the cortex. With continued research, we may be able to map the spatially biased parcellation observed here to that ground truth or find a practical means of acquiring a spatially unbiased BOLD signal. Until then, our findings suggest that ROIs based on common gyral-based labeling systems (e.g., Caviness et al., 1996; Desikan et al., 2006; Lancaster et al., 2000) may be suboptimal for assessing speech-related BOLD effects because they mix responses from distinct functional areas in the core motor, somatosensory, and auditory regions of the brain, resulting in a potential loss of power. Likewise, failing to combine similar responses, for instance, across the precentral sulcus, would also adversely affect power.

The boundary completion step formed fully closed fROIs by following the high-probability contours of the population boundary distribution map. Medial-to-lateral subdivisions of motor, premotor, and somatosensory cortex were formed, and anterior-to-posterior subdivisions were formed along the superior temporal cortex. Overall, the parcellation of the core speech production areas is similar in the two hemispheres, with the notable exception of a more finely parceled superior temporal gyrus and insula in the left hemisphere compared to the right.

Post hoc clustering of the fROIs based on their mean speech response patterns gives us a picture of the functional organization of the speech-responsive cortical areas (see Figure 5). We see a core sensorimotor network of fROIs in the bilateral orofacial somatomotor and auditory cortex that comprises the entirety of the top branch of the clustering dendrogram (see Supplemental Material S2) and is unique in terms of its large response distance from other

observed speech-positive networks. Nearest this network, both in terms of spatial distance on the cortical surface and functional response distance, is a cluster of fROIs in adjacent higher order premotor and auditory cortex, anterior insula, and medial prefrontal cortex that we labeled as the primary flanking network. Together, these networks represent much of the “minimal speech production” network described by Bohland and Guenther (2006) and encompass cortical regions consistently shown to be active during speech production (Basilakos et al., 2017; Brown et al., 2005; Indefrey, 2011; Sörös et al., 2006; Turkeltaub et al., 2002; Wise et al., 1999). The division of this “minimal” network into subnetworks with distinct functional response patterns reflects commonly hypothesized broad roles for these areas in neural models of speech production (e.g., Guenther et al., 2006; Hickok, Houde, & Rong, 2011): a core sensorimotor network involved in speech motor execution (issuing motor commands to the articulators, encoding the consequent sensory feedback) and a flanking network of areas involved in speech motor planning (motor program selection, initiation, and monitoring).

Missing from these two networks is most of the left inferior frontal gyrus, including most of Broca’s area and the inferior frontal sulcus. The former region is classically thought to play an important role in language production (e.g., Heim, Opitz, & Friederici, 2003; Vigneau et al., 2006) and speech motor planning (e.g., Guenther, 2016); the latter has been associated with speech sequence planning (e.g., Bohland & Guenther, 2006) and more generally with verbal and nonverbal working memory (Daniel, Katz, & Robinson, 2016; Rottschy et al., 2012). Rather than clustering with the primary flanking network, fROIs representing these areas clustered with the secondary flanking network described in the Results section. This network also includes fROIs from the junction of orbital and insular cortex and the cingulate gyrus, bilaterally, and a large fROI spanning the posterior lateral occipitotemporal sulcus in the left hemisphere (*l.lots.3* in Figure 5 and Table 2). That fROI includes the “visual word form area,” a left-lateralized brain region critically involved in visual word recognition (Dehaene & Cohen, 2011), suggesting that perhaps this network is involved in mapping the visual representation of the stimulus to a representation that serves as an input to the speech motor system.

Bolstering this notion are studies that have identified intrinsic cortical networks based on resting state functional connectivity. Power et al. (2011) describe a portion of the intrinsic cingulo-opercular network (Dosenbach et al., 2006) that overlaps with the medial (cingulate and adjacent superior frontal), orbital, and insular portions of the secondary flanking network and corresponds to the “salience” network described by Seeley et al. (2007). According to V. Menon (2015), the anterior insula, which receives inputs from auditory and visual cortices, contributes to the salience network by detecting behaviorally relevant stimuli, whereas the dorsal anterior cingulate cortex and adjacent medial prefrontal cortex are more directly involved in response selection and monitoring.

We have previously hypothesized that the left inferior frontal sulcus acts as a phonological working memory buffer, interacting with the medial prefrontal cortex (SMA and preSMA) and the posterior inferior frontal gyrus (Bohland et al., 2010), so it is somewhat surprising that this region clustered with the secondary rather than the primary flanking network. It also did not cluster with the intraparietal and dorsal premotor network, which comprises areas of the cortex that contribute to working memory (Daniel et al., 2016; Rottschy et al., 2012) and largely overlaps with the frontoparietal intrinsic network, another cognitive control or “central executive” network.

Another interesting aspect of the secondary flanking network is the small portion of the inferior frontal gyrus included from the right hemisphere (much of what is excluded from that ROI in the right hemisphere is the inferior frontal sulcus). Most of the right inferior frontal cortex instead clusters with nearby premotor and middle frontal fROIs, an fROI in the right supramarginal gyrus, and another in the right superior temporal sulcus fROI, to form a fully right-lateralized network. Based on a growing body of evidence, we have hypothesized that lateral premotor and inferior frontal cortex in the right hemisphere contributes to response monitoring and feedback-based corrective control of articulation (see Guenther, 2016). We expect this system to be engaged during normal speech production (monitoring), but particularly when a sensory error is detected, which would be expected in some of the experimental conditions in our sample, for example, those that included sensory feedback perturbation or required speakers to produce unfamiliar phonetic sequences such as illegal consonant clusters. The right premotor and middle superior temporal sulcus network noted here is further evidence that the right premotor and inferior frontal cortex is part of a right-lateralized network that, combined with prior results from studies of response monitoring (Fu et al., 2005) and sensory perturbations during speech (Niziolek & Guenther, 2013; Tourville, Reilly, & Guenther, 2008; Toyomura et al., 2007), appears to be involved in feedback-based control of speech.

Future Directions

We consider the cortical fROIs described here to be a starting point rather than an end point in the development of an optimal means of comparing responses across speakers. The fROIs illustrated in Figures 2–5 were formed by the most reliable boundaries derived from our sample. Over the majority of cortex, within-fROI speech responses were uniform, but there were some exceptions (see Figure 4). For instance, the large fROIs that formed across the frontal and central opercular areas in both hemispheres contain multiple, spatially separated peaks in the speech–baseline contrast. These regions may be better segregated by a larger array of experimental tasks or if information from other imaging modalities is taken into account. In the future, additional speaking conditions will be added to the pooled data to further refine the functional parcellation. We also plan to integrate anatomical boundaries into the fROI

parcellation in areas where low reliability of boundary locations prevented an adequate functional parcellation and a clear anatomical marker is available.

On the other hand, our confidence in the fROI boundaries we derived varies; boundaries that align with areas of higher overlap in the population distribution map are more likely to represent a functional transition than those along areas of lower overlap. Some boundaries may warrant removal, depending on the reliability of the boundary and a comparison of response variability within the fROIs that share a border to that of the combined fROI that would result from the boundary, among other factors. In the future, we plan to explore and formalize this process.

Another important future step in the functional parcellation of speech regions of the brain is to expand the current effort to subcortical regions involved in speech production including the basal ganglia, thalamus, and cerebellum. Like cortical parcellation, a functional parcellation of these areas will improve our understanding of their roles in speech processes and offers the potential for improved analysis of speech neuroimaging data. It was also informative to explore how the mean response of fROIs within these subcortical areas cluster with those of the cortical fROIs.

Finally, we have only scratched the surface of what we can learn about the functions represented by the fROIs and the functional networks we observed. We have identified fROIs with relatively low within-ROI variance across a wide range of speech tasks; now, we want to investigate how that response is modulated (presumably uniformly) by specific speaking conditions. Activity from which fROIs/networks correlate with sensory error? Which correlate with syllable complexity? Are there regions that can predict phonemic content? Our pooled analysis platform allows us to address such questions. For instance, there are distinct somatomotor and auditory subclusters in the core sensorimotor network (see Supplemental Material S2), but region *Lstg.4* branches off from the other auditory regions, with a shorter response distance from the somatomotor fROIs of the network than the other auditory fROIs. By characterizing the activity in this area across a wide range of speaking conditions, we hope to differentiate this area from the other auditory cortices and gain a better understating of its contribution to the neural control of speech.

Acknowledgments

This research was supported by grants from the National Institute on Deafness and Other Communication Disorders: R01 DC007683 and DC002852 (principal investigator: F. G.). The content of this report is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

Basilakos, A., Smith, K. G., Fillmore, P., Fridriksson, J., & Fedorenko, E. (2017). Functional characterization of the human speech articulation network. *Cerebral Cortex*, 28, 1816–1830.

- Beal, D. S., Segawa, J. A., Tourville, J. A., Cai, S., & Guenther, F. H. (2012). *Speech motor sequence learning difficulties in persistent developmental stuttering: An fMRI study* (Program No. 681.0). Meeting Planner, 42nd Annual Meeting of the Society for Neuroscience, New Orleans, LA.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57, 289–300.
- Bohland, J. W., Bullock, D., & Guenther, F. H. (2010). Neural representations and mechanisms for the performance of simple speech sequences. *Journal of Cognitive Neuroscience*, 22(7), 1504–1529.
- Bohland, J. W., & Guenther, F. H. (2006). An fMRI investigation of syllable sequence production. *NeuroImage*, 32(2), 821–841.
- Brown, S., Ingham, R. J., Ingham, J. C., Laird, A. R., & Fox, P. T. (2005). Stuttered and fluent speech production: An ALE meta-analysis of functional neuroimaging studies. *Human Brain Mapping*, 25(1), 105–117.
- Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, 1124(1), 1–38.
- Cai, S., Tourville, J. A., Beal, D. S., Perckell, J. S., Guenther, F. H., & Ghosh, S. S. (2014). Diffusion imaging of cerebral white matter in persons who stutter: Evidence for network-level anomalies. *Frontiers in Human Neuroscience*, 8, 54.
- Caviness, V. S., Jr., Meyer, J., Makris, N., & Kennedy, D. N. (1996). MRI-based topographic parcellation of human neocortex: An anatomically specified method with estimate of reliability. *Journal of Cognitive Neuroscience*, 8(6), 566–587.
- Cohen, A. L., Fair, D. A., Dosenbach, N. U., Miezin, F. M., Dierker, D., Van Essen, D. C., . . . Petersen, S. E. (2008). Defining functional areas in individual human brains using resting functional connectivity MRI. *NeuroImage*, 41(1), 45–57.
- Costafreda, S. G. (2009). Pooling fMRI data: Meta-analysis, mega-analysis and multi-center studies. *Frontiers in Neuroinformatics*, 3, 33.
- Daniel, T. A., Katz, J. S., & Robinson, J. L. (2016). Delayed match-to-sample in working memory: A BrainMap meta-analysis. *Biological Psychology*, 120, 10–20.
- Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, 15(6), 254–262.
- Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., . . . Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, 31(3), 968–980.
- Dosenbach, N. U., Visscher, K. M., Palmer, E. D., Miezin, F. M., Wenger, K. K., Kang, H. C., . . . Petersen, S. E. (2006). A core system for the implementation of task sets. *Neuron*, 50(5), 799–812.
- Duvernoy, H. M. (1999). *The human brain: Surface, blood supply, and three dimensional Anatomy* (2nd ed.). New York, NY: Springer-Verlag.
- Efron, B. (1982). *The jackknife, the bootstrap, and other resampling plans*. CBMS-NSF Regional Conference Series in Applied Mathematics, Monograph 38, Philadelphia, PA.
- Eickhoff, S. B., Bzdok, D., Laird, A. R., Roski, C., Caspers, S., Zilles, K., & Fox, P. T. (2011). Co-activation patterns distinguish cortical modules, their connectivity and functional differentiation. *NeuroImage*, 57(3), 938–949.
- Eickhoff, S. B., Heim, S., Zilles, K., & Amunts, K. (2009). A systems perspective on the effective connectivity of overt speech production. *Philosophical Transactions of the Royal Society of London*

- A: *Mathematical, Physical and Engineering Sciences*, 367(1896), 2399–2421.
- Fischl, B. (2012). FreeSurfer. *NeuroImage*, 62(2), 774–781.
- Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, 9(2), 195–207.
- Fu, C. H., Vythelingum, G. N., Brammer, M. J., Williams, S. C., Amaro, E., Jr., Andrew, C. M., ... McGuire, P. K. (2005). An fMRI study of verbal self-monitoring: Neural correlates of auditory verbal feedback. *Cerebral Cortex*, 16(7), 969–977.
- Gagnon, L., Sakadžić, S., Lesage, F., Musacchia, J. J., Lefebvre, J., Fang, Q., ... Boas, D. A. (2015). Quantifying the microvascular origin of BOLD-fMRI from first principles with two-photon microscopy and an oxygen-sensitive nanoprobe. *The Journal of Neuroscience*, 35(8), 3663–3675.
- 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. *Journal of Speech, Language, and Hearing Research*, 51(5), 1183–1202.
- Golfopoulos, E., & Guenther, F. H. (2011). *Prominence in English spoken utterances: fMRI evidence for left hemisphere cortical recruitment*. Poster presented at the 17th Annual Meeting of the Organization on Human Brain Mapping, Québec City, Canada.
- Golfopoulos, E., Tourville, J. A., Bohland, J. W., Ghosh, S. S., Nieto-Castañón, A., & Guenther, F. H. (2011). fMRI investigation of unexpected somatosensory feedback perturbation during speech. *NeuroImage*, 55(3), 1324–1338.
- Guenther, F. H. (2016). *Neural control of speech*. Cambridge, MA: MIT Press.
- Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language*, 96(3), 280–301.
- Hagler, D. J., Jr., Saygin, A. P., & Sereno, M. I. (2006). Smoothing and cluster thresholding for cortical surface-based group analysis of fMRI data. *NeuroImage*, 33(4), 1093–1103.
- Hayasaka, S., & Nichols, T. E. (2003). Validating cluster size inference: Random field and permutation methods. *NeuroImage*, 20(4), 2343–2356.
- Heim, S., Opitz, B., & Friederici, A. D. (2003). Distributed cortical networks for syntax processing: Broca's area as the common denominator. *Brain and Language*, 85(3), 402–408.
- Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: Computational basis and neural organization. *Neuron*, 69(3), 407–422.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8(5), 393–402.
- Hinds, O. P., Rajendran, N., Polimeni, J. R., Augustinack, J. C., Wiggins, G., Wald, L. L., ... Fischl, B. (2008). Accurate prediction of V1 location from cortical folds in a surface coordinate system. *NeuroImage*, 39(4), 1585–1599.
- Houde, J. F., & Nagarajan, S. S. (2011). Speech production as state feedback control. *Frontiers in Human Neuroscience*, 5, 82.
- Indefrey, P. (2011). The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology*, 2, 255.
- Kim, J.-H., Lee, J.-M., Jo, H. J., Kim, S. H., Lee, J. H., Kim, S. T., ... Ziad, S. S. (2010). Defining functional SMA and pre-SMA subregions in human MFC using resting state fMRI: Functional connectivity-based parcellation method. *NeuroImage*, 49(3), 2375–2386.
- Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., ... Fox, P. T. (2000). Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping*, 10(3), 120–131.
- Lee, M. H., Smyser, C. D., & Shimony, J. S. (2013). Resting-state fMRI: A review of methods and clinical applications. *American Journal of Neuroradiology*, 34(10), 1866–1872.
- Menon, R. S. (2012). The great brain versus vein debate. *NeuroImage*, 62(2), 970–974.
- Menon, V. (2015). Salience network. In A. W. Toga (Ed.), *Brain mapping: An encyclopedic reference* (Vol. 2, pp. 597–611). Cambridge, MA: Academic Press.
- Nieto-Castañón, A., Ghosh, S. S., Tourville, J. A., & Guenther, F. H. (2003). Region of interest based analysis of functional imaging data. *NeuroImage*, 19(4), 1303–1316.
- Niziolek, C. A., & Guenther, F. H. (2013). Vowel category boundaries enhance cortical and behavioral responses to speech feedback alterations. *The Journal of Neuroscience*, 33(29), 12090–12098.
- Overduin, S. A., & Guenther, F. H. (2009). *Brain structures differentially responsible for controlling overt and covert speech* [Abstract]. Chicago, IL: Society for Neuroscience.
- Park, H., Park, Y.-H., Cha, J., Seo, S. W., Na, D. L., & Lee, J.-M. (2017). Agreement between functional connectivity and cortical thickness-driven correlation maps of the medial frontal cortex. *PLOS ONE*, 12(3), e0171803.
- Peeva, M. G., Guenther, F. H., Tourville, J. A., Nieto-Castañón, A., Anton, J.-L., Nazarian, B., & Alario, F.-X. (2010). Distinct representations of phonemes, syllables, and supra-syllabic sequences in the speech production network. *NeuroImage*, 50(2), 626–638.
- Petrides, M. (2014). *Neuroanatomy of language regions of the human brain*. London, United Kingdom: Academic Press.
- Poldrack, R. A. (2007). Region of interest analysis for fMRI. *Social Cognitive and Affective Neuroscience*, 2(1), 67–70.
- Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., ... Petersen, S. E. (2011). Functional network organization of the human brain. *Neuron*, 72(4), 665–678.
- Rademacher, J., Galaburda, A. M., Kennedy, D. N., Filipek, P. A., & Caviness, V. S., Jr. (1992). Human cerebral cortex: Localization, parcellation, and morphometry with magnetic resonance imaging. *Journal of Cognitive Neuroscience*, 4(4), 352–374.
- Rottschy, C., Langner, R., Dögan, I., Reetz, K., Laird, A. R., Schulz, J. B., ... Eickhoff, S. B. (2012). Modelling neural correlates of working memory: A coordinate-based meta-analysis. *NeuroImage*, 60(1), 830–846.
- Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., ... Greicius, M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *The Journal of Neuroscience*, 27(9), 2349–2356.
- Segawa, J. A., Tourville, J. A., Beal, D. S., & Guenther, F. H. (2013). *The representation of syllabic frame structures and phonological content in the brain*. Poster presented at the 19th Annual Meeting of the Organization for Human Brain Mapping, Seattle, WA.
- Segawa, J. A., Tourville, J. A., Beal, D. S., & Guenther, F. H. (2015). The neural correlates of speech motor sequence learning. *The Journal of Cognitive Neuroscience*, 27(4), 819–831.
- Seghier, M. L., & Price, C. J. (2009). Dissociating functional brain networks by decoding the between-subject variability. *NeuroImage*, 45(2), 349–359.
- Simonyan, K., Ackermann, H., Chang, E. F., & Greenlee, J. D. (2016). New developments in understanding the complexity of human speech production. *The Journal of Neuroscience*, 36(45), 11440–11448.
- Sörös, P., Sokoloff, L. G., Bose, A., McIntosh, A. R., Graham, S. J., & Stuss, D. T. (2006). Clustered functional MRI of overt speech production. *NeuroImage*, 32(1), 376–387.
- Takai, O., Brown, S., & Liotti, M. (2010). Representation of the speech effectors in the human motor cortex: Somatotopy or overlap. *Brain and Language*, 113(1), 39–44.

- Talairach, J., & Tournoux, P.** (1988). *Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: An approach to cerebral imaging*. New York, NY: Thieme.
- Thirion, B., Varoquaux, G., Dohmatob, E., & Poline, J.-B.** (2014). Which fMRI clustering gives good brain parcellations. *Frontiers in Neuroscience*, 8, 167.
- Tourville, J., & Guenther, F. H.** (2012). *Automatic cortical labeling system for neuroimaging studies of normal and disordered speech* (Program No. 681.06). Meeting Planner, 42nd Annual Meeting of the Society for Neuroscience, New Orleans, LA.
- Tourville, J. A., & Guenther, F. H.** (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes*, 26(7), 952–981.
- Tourville, J. A., Reilly, K. J., & Guenther, F. H.** (2008). Neural mechanisms underlying auditory feedback control of speech. *NeuroImage*, 39(3), 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(2), 499–503.
- Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A.** (2002). Meta-analysis of the functional neuroanatomy of single-word reading: Method and validation. *NeuroImage*, 16(3), 765–780.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., . . . Joliot, M.** (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, 15(1), 273–289.
- Van Essen, D. C., Drury, H. A., Joshi, S., & Miller, M. I.** (1998). Functional and structural mapping of human cerebral cortex: Solutions are in the surfaces. *Proceedings of the National Academy of Sciences of the United States of America*, 95(3), 788–795.
- Viessmann, O. M., Bianciardi, M., Scheffler, K., Wald, L. L., & Polimeni, J. R.** (2018). *The EPI rs-fMRI signal shows an orientation effect with respect to B0 and phase-encode axis across cortical depth*. Joint Annual Meeting ISMRM-ESMRMB 2018, Paris, France.
- Vigneau, M., Beaucousin, V., Hervé, P.-Y., Duffau, H., Crivello, F., Houdé, O., . . . Tzourio-Mazoyer, N.** (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *NeuroImage*, 30(4), 1414–1432.
- Ward, J. H., Jr.** (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301), 236–244.
- Whitfield-Gabrieli, S., & Nieto-Castanón, A.** (2012). Conn: A functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connectivity*, 2(3), 125–141.
- Wilson, S. M.** (2014). The impact of vascular factors on language localization in the superior temporal sulcus. *Human Brain Mapping*, 35(8), 4049–4063.
- Wise, R. J., Greene, J., Büchel, C., & Scott, S. K.** (1999). Brain regions involved in articulation. *The Lancet*, 353(9158), 1057–1061.
- Yousry, T., Schmid, U. D., Alkadhi, H., Schmidt, D., Peraud, A., Buettner, A., & Winkler, P.** (1997). Localization of the motor hand area to a knob on the precentral gyrus: A new landmark. *Brain*, 120(1), 141–157.
- Zilles, K., & Amunts, K.** (2010). Centenary of Brodmann's map—Conception and fate. *Nature Reviews Neuroscience*, 11, 139–145.