

ORIGINAL ARTICLE

Functional Characterization of the Human Speech Articulation Network

Alexandra Basilakos¹, Kimberly G. Smith^{1,2}, Paul Fillmore³, Julius Fridriksson¹ and Evelina Fedorenko^{4,5,6}

¹Department of Communication Sciences and Disorders, University of South Carolina, Columbia, SC 29208, USA, ²Department of Speech Pathology and Audiology, University of South Alabama, Mobile, AL 36688, USA,

³Department of Communication Sciences and Disorders, Baylor University, Waco, TX 76798, USA,

⁴Department of Psychiatry, Harvard Medical School, Boston, MA 02115, USA, ⁵Department of Psychiatry, Massachusetts General Hospital, Boston, MA 02114, USA and ⁶Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Address correspondence to Alexandra Basilakos, University of South Carolina, Columbia, SC 29208, USA. Email: basilaka@mailbox.sc.edu; Evelina Fedorenko, Massachusetts General Hospital, Boston, MA 02114, USA. Email: evelina.fedorenko@mgh.harvard.edu

Abstract

A number of brain regions have been implicated in articulation, but their precise computations remain debated. Using functional magnetic resonance imaging, we examine the degree of functional specificity of articulation-responsive brain regions to constrain hypotheses about their contributions to speech production. We find that articulation-responsive regions (1) are sensitive to articulatory complexity, but (2) are largely nonoverlapping with nearby domain-general regions that support diverse goal-directed behaviors. Furthermore, premotor articulation regions show selectivity for speech production over some related tasks (respiration control), but not others (nonspeech oral-motor [NSO] movements). This overlap between speech and nonspeech movements concurs with electrocorticographic evidence that these regions encode articulators and their states, and with patient evidence whereby articulatory deficits are often accompanied by oral-motor deficits. In contrast, the superior temporal regions show strong selectivity for articulation relative to nonspeech movements, suggesting that these regions play a specific role in speech planning/production. Finally, articulation-responsive portions of posterior inferior frontal gyrus show some selectivity for articulation, in line with the hypothesis that this region prepares an articulatory code that is passed to the premotor cortex. Taken together, these results inform the architecture of the human articulation system.

Key words: articulation, executive functions, fMRI, speech production

Introduction

Fluent speech—a remarkable, uniquely human behavior—requires the planning of sound sequences, followed by the execution of corresponding motor plans. In addition, articulation depends upon respiratory support for phonation and rapid oral-motor movements to shape airflow into speech sounds. Although extensive theorizing and a wealth of experimental data have yielded sophisticated models of articulation (e.g., [Browman and Goldstein 1995](#); [Guenther 2006](#); [Perkell 2012](#); [Buchwald 2014](#);

[Dell 2014](#); [Goldrick 2014](#); [Ziegler and Ackermann 2014](#); [Blumstein and Baum 2016](#)), our understanding of how these processes are implemented in neural tissue remains limited.

Much of the early evidence implicating particular brain regions in speech production comes from investigations of patients with apraxia of speech (AOS), a motor speech disorder that most commonly results from stroke-induced damage to the dominant (typically left) hemisphere (e.g., [Darley et al. 1975](#); [Dronkers 1996](#); [Dronkers and Ogar 2004](#); [Hillis et al. 2004](#);

Duffy 2005; Ogar et al. 2006; Richardson et al. 2012). Speech production in AOS is characterized by sound distortions, distorted substitutions and/or additions, visible and/or audible articulatory groping, frequent but often failed attempts at self-correction, and atypical prosody (Darley et al. 1975; Rosenbek et al. 1984; Ogar et al. 2005; Strand et al. 2014). A number of cortical regions, mostly in the frontal lobe, have been linked to AOS, including the left inferior frontal gyrus and surrounding areas (e.g., Hillis et al. 2004; Richardson et al. 2012), the anterior insula (e.g., Dronkers 1996; Dronkers et al. 2004; Baldo et al. 2011), premotor and supplementary motor regions (e.g., Jonas 1981; Pai 1999; Josephs et al. 2006, 2012; Whitwell et al. 2013; Graff-Radford et al. 2014; Basilakos et al. 2015; Itabashi et al. 2016), and parts of the postcentral gyrus (e.g., Hickok et al. 2014; Basilakos et al. 2015). Less frequent are reports of AOS following infarct to the basal ganglia (e.g., Kertesz 1984) or the right cerebellum (see e.g., Mariën et al. 2015 and Ziegler 2016 for discussions of the role of the cerebellum in articulation).

Neuroimaging studies have implicated a similar set of regions, as well as additional regions in the superior temporal cortices, in typical speakers (e.g., Wise et al. 1999; Bohland and Guenther 2006; Eickhoff et al. 2009; Fedorenko et al. 2015). This convergence on the same candidate “articulation network”—a set of brain regions that presumably work together to achieve the complex task of speech production (e.g., Mesulam 1998)—is encouraging. However, the computations that these brain regions support, and the division of labor among them, remain debated (e.g., Blumstein and Baum 2016), with different proposals assigning different roles to some of the same brain regions (e.g., Hickok 2014; Guenther 2016).

One way to constrain hypotheses about the possible computations of a brain region is to characterize its functional response profile: What manipulations is it sensitive to? How selective is its response? What are the necessary and sufficient conditions under which it gets engaged? We here report a functional magnetic resonance imaging (fMRI) study that functionally characterizes the articulation-responsive brain regions, with a focus on the degree of their *functional selectivity*.

Although a number of prior fMRI investigations (e.g., Wildgruber et al. 1996; Wise et al. 1999; Bohland and Guenther 2006; Sörös et al. 2006; Takai et al. 2010) have examined similar questions, they have almost exclusively relied on the traditional group-averaging approach, where individual brains are aligned in the common brain space and voxel-wise functional correspondence is assumed. This approach makes it difficult to compare results across individuals and studies. In particular, to do so, researchers often resort to discussing the observed activations at the level of macroanatomic landmarks (gyri and sulci), which is problematic given the high anatomical and functional heterogeneity that characterizes these large structures (e.g., Amunts et al. 2010; Fedorenko et al. 2012; Deen et al. 2015), each of which is comprised of many cubic centimeters of cortex. Thus, observing activation for two different manipulations within, say, the inferior frontal gyrus does not afford the inference that these manipulations rely on the same neural mechanism (see also Poldrack 2011; Yarkoni et al. 2011). To circumvent these difficulties, we adopt a functional localization approach (e.g., Saxe et al. 2006; Fedorenko et al. 2010), where we first identify—in each participant individually—regions that are active during an articulation task, and then probe these functionally defined regions of interest for their (1) sensitivity to articulatory complexity, and, critically, (2) selectivity relative to tasks that share some features with articulation and/or have been previously shown to produce responses within the same macroanatomical brain areas. In addition to its higher sensitivity and functional resolution

(Nieto-Castañón and Fedorenko 2012), this approach provides a straightforward way to pool data across individuals and manipulations, and can be used in future studies to compare any new manipulation to the ones examined here.

In this study, we asked 3 key research questions: (1) Are articulation-responsive regions sensitive to articulatory complexity, as would be expected based on prior studies (e.g., Shuster and Lemieux 2005; Bohland and Guenther 2006; Papoutsi et al. 2009)?; (2) To what extent are articulation-responsive regions functionally selective for speech production, relative to ancillary behaviors, like respiration and nonspeech oral-motor movements?; and (3) What is the relationship between articulation-responsive regions and the domain-general multiple demand (MD) system implicated in executive functions (like attention and cognitive control) and active during diverse goal-directed behaviors (Duncan 2010, 2013)? The latter question is important given that (1) articulation has been shown to elicit neural responses within some of the macroanatomical brain regions linked to general cognitive demand, including the inferior frontal gyrus, the anterior insula, the precentral gyrus, and the supplementary motor area (SMA), and (2) articulation is a demanding task (perhaps especially so for unfamiliar pseudowords; e.g., Segawa et al. 2015), and the regions of the MD system have been shown to respond to demanding tasks across domains (Duncan and Owen 2000; Fedorenko et al. 2013; Hugdahl et al. 2015). Thus, some of the brain regions previously discussed as belonging to the articulation network may contribute to articulation via highly domain-general processes, and this would fundamentally alter the hypothesized role of these regions in the existing neural models of articulation. We talk about how our results inform existing models in the Discussion.

Methods

Participants

Twenty healthy adults (19 females; mean age = 26 years, range 18–51) were recruited from the University of South Carolina community. All participants were right-handed native speakers of English. No participants reported history of neurological, speech, language, or hearing problems. The study was approved by the University of South Carolina Institutional Review Board. All participants gave written informed consent for study inclusion and were given course credit or payment for participation.

Design, Materials, and Procedure

Each participant performed 3 experiments: (1) the critical articulation experiment, (2) an experiment used to evaluate the functional selectivity of the articulation-responsive brain regions, and (3) an experiment used to evaluate potential overlap with spatially proximal domain-general brain regions implicated in attention and cognitive control.

In Experiments 1 and 2, participants viewed 2500-ms-long video clips of a female actor (native English speaker) and were asked to imitate each action (during the subsequent 2500 ms) in a blocked design. The clips were created so that the duration was as close to 2500 ms as possible, and some behavioral piloting of the tasks suggested that participants generally follow the timing closely in their imitations. Each block included 4 trials and lasted 20 s. The sound was delivered via Resonance Technology’s magnetic resonance imaging (MRI)-compatible Serene Sound headphones.

In Experiment 1, 192 bisyllabic pseudowords were used: The second syllable was always a simple consonant–vowel (CV) syllable (e.g., /fʌ, si, gɒ/), and the first syllable varied between the hard and the easy condition. The hard condition included

phonotactically legal consonant clusters, including many that require fast transitions between articulator positions, and took the form of CCVCC syllables (e.g., /krou:rd, dreilf, sfi:lt/). The easy condition included no clusters and took the form of CV syllables (e.g., /mo:, koi, rei/). For example, sample items in the hard condition sounded like /snoirb-ga/ or /kweips-ki/, and sample items in the easy condition sounded like /lo:-ka/ or /guma/. Pseudowords were used instead of real words in order to isolate the process of speech production (articulation) from the higher level components of language processing (i.e., lexico-semantic processes). The use of pseudowords that obey the phonotactic constraints of English ensures that the articulatory processes required to produce these are similar to those used for producing real words (e.g., Hickok and Poeppel 2004; Fedorenko and Thompson-Schill 2014). Participants completed two 460-s-long runs, each containing 9 blocks per condition and 5 fixation blocks. Any given participant was exposed to 72 hard and 72 easy pseudowords, which were sampled from the larger set of 96 hard and 96 easy pseudowords (a complete list of materials, including the video clips, is available at <https://evlab.mit.edu/papers/artic>). During the fixation block, a cross appeared in the center of the screen, and participants were asked to look at the screen and rest. Participants were told that these blocks were used to measure baseline brain activity.

In Experiment 2, 10 tokens each were used of (1) vowels (e.g., /i:/ as in feet; Fig. 1a), (2) respiration sequences (e.g., inhale slowly then exhale quickly), and (3) nonspeech oral-motor movements (e.g., pucker of the lips; Fig. 1b). Each trial included a single vowel, respiration sequence, or nonspeech oral-motor movement. The respiration and nonspeech oral-motor movements conditions did not involve phonation. Participants completed three 480-s-long runs, each containing 6 blocks per condition and 6 fixation blocks.

Finally, we included an executive (spatial working memory [WM]) task (Experiment 3; adapted from Fedorenko et al. 2013) to evaluate potential overlap between articulation-responsive regions and spatially proximal domain-general brain regions implicated in attention and cognitive control (e.g., Duncan 2010, 2013; Fedorenko et al. 2013). These regions have been shown to be sensitive to effort across diverse tasks (Duncan and Owen 2000; Hugdahl et al. 2015) and can thus be defined using any task that manipulates difficulty, by contrasting activation for a harder versus an easier condition (Fedorenko et al. 2013). In the task we used here (see also Blank et al. 2014),

participants saw a 3×4 grid and kept track of 4 or 8 locations in the easy and hard condition, respectively, in a blocked design. At the end of each trial, participants had to choose the grid with the correct locations in a 2-alternative forced-choice question (Fig. 1c). Each block included four 8.5 s trials and lasted 34 s (see Fedorenko et al. 2013 for details). Participants completed two 404-s-long runs, each containing 5 blocks per condition and five 16-s-long fixation blocks.

Condition order was palindromic within each run (to avoid order effects) and counterbalanced across runs and participants for all experiments, and the order of trials was randomized across participants. All participants practiced each task outside of the scanner. They were instructed to follow the timing of the actions as closely as possible, ensuring that timing was similar across all participants and scanning sessions. The entire scanning session lasted approximately 1.5 h.

Image Acquisition

Structural and functional MRI data were collected using a whole-body 3 T Siemens Trio scanner equipped with a 12-element head coil. T1-weighted structural images were collected in 192 axial slices with 1 mm isotropic voxels (time repetition [TR] = 2250 ms and time echo [TE] = 4.2 ms). Functional, blood oxygen level dependent (BOLD) data were acquired using an EPI sequence (with a 76° flip angle) with the following parameters: thirty-one 3.75 mm-thick near-axial slices acquired in sequential order (with 25% distance factor), 2 mm \times 2 mm in-plane resolution, FoV = 208 \times 208 mm, A > P phase encoding, TR = 2000 ms and TE = 30 ms. The first 8 s of each run were excluded to allow for steady state magnetization.

Image Preprocessing

MRI data were analyzed using SPM5 and the *spm_ss* toolbox designed to perform second-level analyses that incorporate information from individual activation maps (available for download at http://www.nitrc.org/projects/spm_ss). Each participant's data were motion-corrected, normalized to the MNI space (Montreal Neurological Institute template), re-sampled into 2 mm isotropic voxels, smoothed using a 4 mm Gaussian kernel, and high-pass filtered (at 200 s). For all tasks, effects were estimated using a General Linear Model in which each experimental condition was modeled with a boxcar function convolved with the canonical hemodynamic response function. The boxcar function modeled entire blocks.

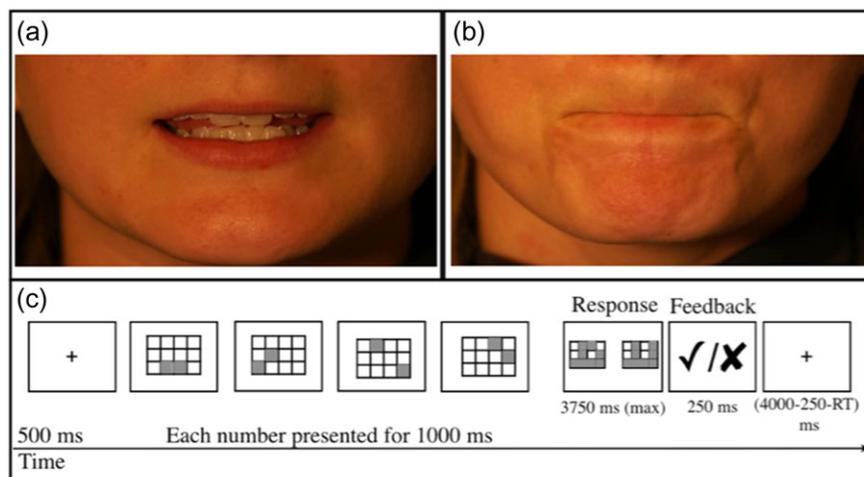


Figure 1. Sample trials from Experiments 2 and 3. (a) Example vowel. (b) Example nonspeech oral-motor movement. (c) A hard trial from the spatial working memory task.

Discovery of Articulation-Responsive Regions

To identify brain regions sensitive to articulation, we performed a group-constrained subject-specific (GSS) analysis (Fedorenko et al. 2010; Julian et al. 2012). To ensure that we did not miss any relevant functional regions, we chose to use the most inclusive and robust contrast from the articulation task: The hard articulation condition (imitation of complex pseudowords) relative to the low-level (fixation) baseline. We took individual whole-brain activation maps for this contrast and binarized them so that voxels that show a reliable response (significant at $P < 0.001$, uncorrected at the whole-brain level) were turned into 1's and all other voxels were turned into 0's. We then overlaid these maps to create a probabilistic activation overlap map (Fig. 2a), thresholded this map to only include voxels where at least 3 of the 20 participants showed activation, and divided it into "parcels" using a watershed image parcellation algorithm (see Fedorenko et al. 2010 for details). Finally, we identified parcels that—when intersected with the individual activation maps—contained supra-threshold (i.e., significant for our contrast of interest at $P < 0.001$, uncorrected) voxels in at least 16 of the 20 (i.e., 80%) individual participants. Fifteen parcels satisfied this criterion. However, 4 of these were located in the occipital cortex (to be expected given that the task involves visual presentation of the to-be-imitated stimuli) and were thus excluded.

Definition of Individual Articulation-Responsive Functional Regions of Interest and Estimating Their Responses

In each of the 11 regions, we estimated the response magnitude to the conditions of Experiment 1 (i.e., hard and easy articulation) in individual participants using an across-runs cross-validation procedure (e.g., Nieto-Castañón and Fedorenko 2012), so that the data used to define the Functional Regions of Interest (fROIs) and

to estimate the responses were independent (e.g., Kriegeskorte et al. 2009; Vul et al. 2009). In particular, each parcel was intersected with each participant's activation map for the *hard articulation > fixation* contrast for the first run of the data. The voxels within the parcel were sorted—for each participant—based on their t-values, and the top 10% of voxels were selected as that participant's fROI. The responses were then estimated using the second run's data. The procedure was repeated using the second run to define the fROIs and the first run to estimate the responses. Finally, the responses were averaged across the left-out runs to derive a single response magnitude per participant per region per condition.

To estimate the responses to the conditions of Experiment 2 (imitation of vowels, respiration sequences, and nonspeech oral-motor movements) and Experiment 3 (hard and easy spatial WM conditions), all the data from Experiment 1 were used for defining the individual fROIs. Statistical tests were performed on these extracted percent BOLD signal change values.

Definition of individual domain-general MD system fROIs and estimating their responses.

In addition to the articulation-responsive fROIs, we defined a set of fROIs for the domain-general MD system (Duncan 2010, 2013), based on a spatial WM task (data from Experiment 3). The regions of this system include parts of the inferior frontal and middle frontal gyri, precentral gyrus, SMA, parietal cortices, insula, and anterior cingulate, and have been implicated in a wide range of goal-directed behaviors. A subset of these regions—those within the inferior frontal gyrus, the precentral gyrus, the SMA, and the insular cortex—appear to land in close proximity to the regions previously implicated in articulation (e.g., Wise et al. 1999; Bohland and Guenther 2006; Eickhoff et al. 2009; Fedorenko et al. 2015). Given that articulation is a demanding task, we wanted to examine potential overlap between these

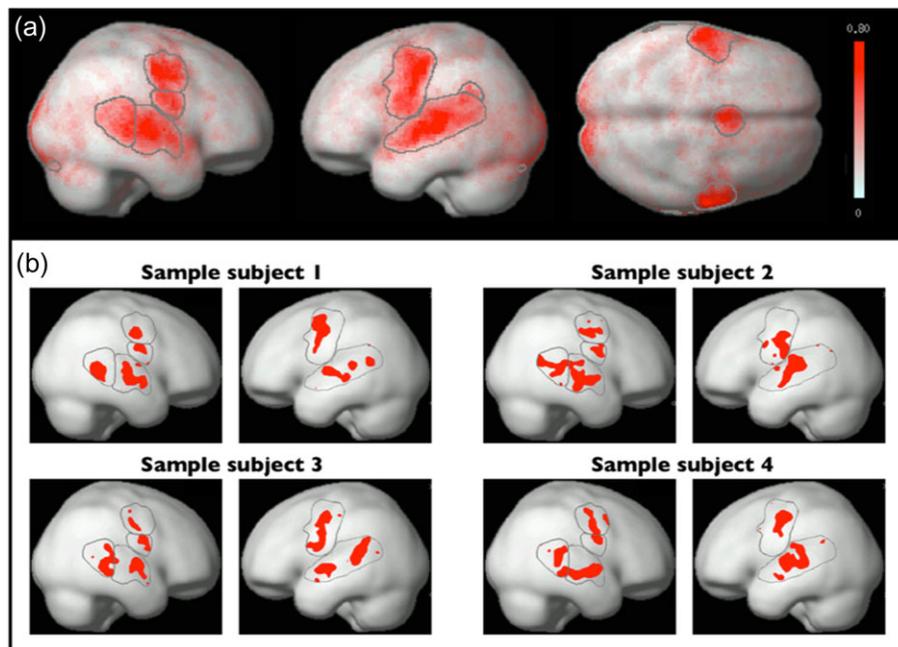


Figure 2. Articulation-responsive brain regions. (a) A whole-brain probabilistic activation overlap map for 20 participants, with the parcels discovered by the GSS analysis (Fedorenko et al. 2010) overlaid on top (note that 1 of the 11 parcels—L pSTG2—is not visible on the surface; see Fig. 3). (b) fROIs in the left and right hemisphere of 4 sample individual participants. These fROIs were defined by intersecting the individual activation maps for the “hard articulation > fixation” contrast with the parcels, and selecting the top 10% of most responsive voxels within each parcel.

domain-general regions and the articulation-responsive regions. Following Fedorenko et al. (2013), we used a set of anatomical parcels from the AAL atlas (Tzourio-Mazoyer et al. 2002) to constrain the selection of individual fROIs. Each of these parcels was intersected with each participant's activation map for the *hard spatial working memory > easy spatial working memory* contrast (e.g., Blank et al. 2014), and the top 10% of voxels were selected as that participant's fROI. To estimate the response magnitude to the spatial WM conditions, we used an across-runs cross-validation procedure, as described above. To estimate the response magnitude to the conditions of Experiments 1 and 2, all the data from Experiment 3 were used for defining the individual fROIs.

Results

The articulation-responsive regions that emerged in the whole-brain GSS analysis (Fedorenko et al. 2010) are shown in Figure 3

(see Fig. 2b, for sample individual fROIs) and include 3 regions in the left superior temporal gyrus (a large region spanning a significant portion of the left superior temporal gyrus—L STG, and 2 small regions in the posterior-most extent of STG—LpSTG1 and LpSTG2), 2 regions in the right superior temporal gyrus including its anterior portion (R aSTG) and its posterior portion (R pSTG), a region in the left precentral gyrus (L PrCG), 2 regions in the right precentral gyrus including its inferior portion (R iPrCG) and its superior portion (R sPrCG), a region in the SMA, and 2 cerebellar regions (LCereb and RCereb). These regions match closely a set of regions implicated in speech articulation in prior patient (e.g., Hillis et al. 2004; Josephs et al. 2006, 2012; Richardson et al. 2012; Whitwell et al. 2013; Basilakos et al. 2015; Itabashi et al. 2016) and brain imaging (e.g., Wise et al. 1999; Bohland and Guenther 2006; Eickhoff et al. 2009; Fedorenko et al. 2015) studies. Thus, our procedure for identifying the key search spaces for defining the individual fROIs was effective, and we can proceed to characterize these regions' response profiles.

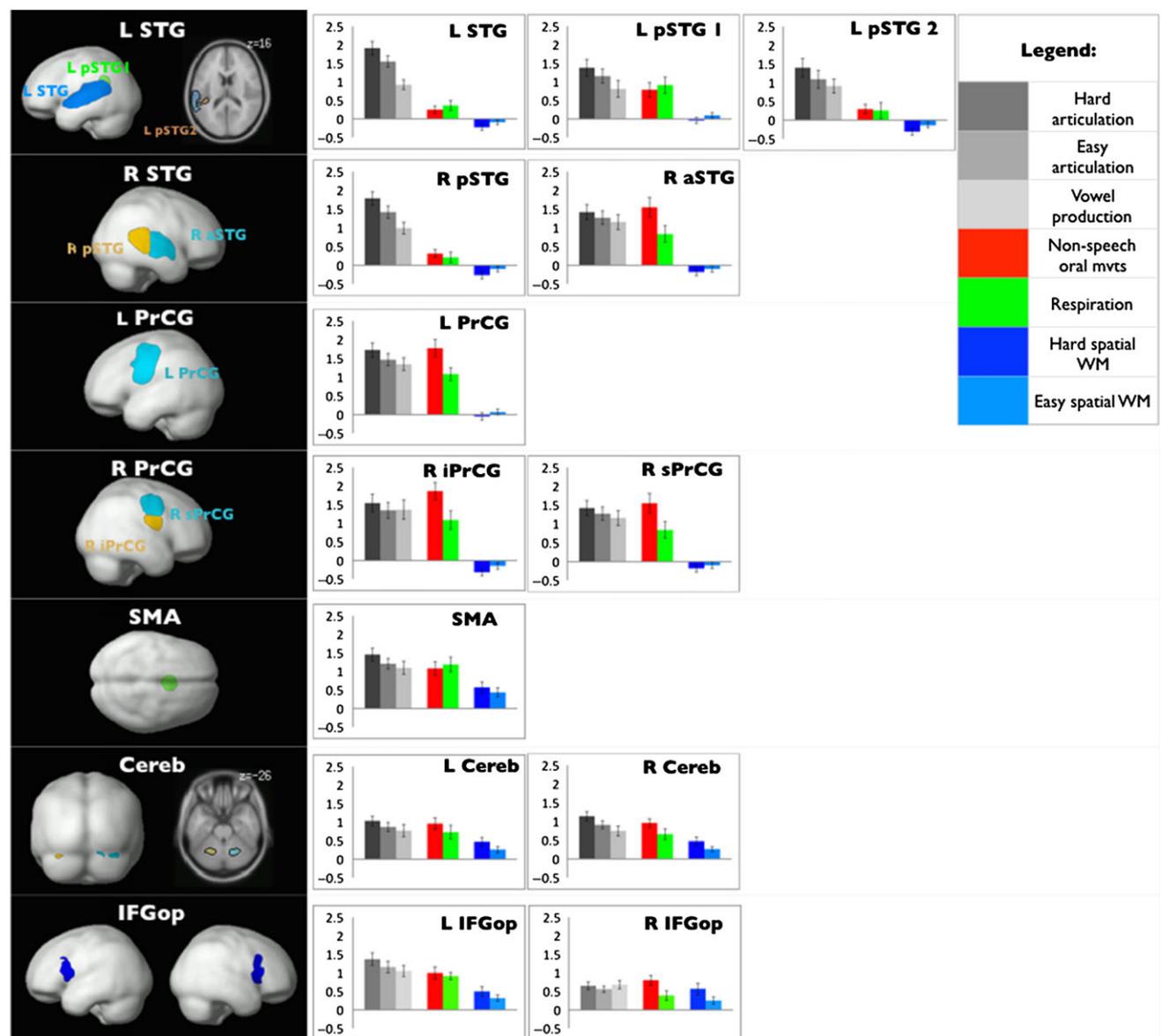


Figure 3. Responses to the articulation and spatial WM conditions of each articulation-responsive fROI. Left: The parcels used to define articulation-responsive fROIs. Right: Responses of the individually defined articulation-responsive fROIs to each condition across Experiments 1–3.

The only brain region—much discussed in the literature as playing a role in articulation (e.g., Broca 1861; Hillis et al. 2004; Richardson et al. 2012) and yet somewhat elusive (e.g., Wise et al. 1999; Flinker et al. 2015)—that did not emerge in our GSS analysis is a region in posterior IFG. This part of the cortex is notoriously variable across individuals (e.g., Amunts et al. 1999; Tomaiuolo et al. 1999; Juch et al. 2005), which plausibly contributes to the elusiveness of this region in previous studies that have assumed voxel-by-voxel correspondence across individuals. In line with this variability, the GSS analysis did discover a small region in the left IFG, but this region was only present in 11 of the 20 participants when intersected with individual activation maps, and hence did not pass our threshold for overlap with at least 16 of the 20 participants. However, given the existing evidence linking this region to articulation, combined with the fact that this region is robustly detectable in individual participants (e.g., Flinker et al. 2015; Long et al. 2016), we include this region in our set, and we recommend that future studies adopt a similar approach, including critically using functional localization, given that posterior IFG contains a number of functionally distinct regions including articulation-responsive regions and domain-general MD regions that we focus on here, but also high-level language processing regions (e.g., Fedorenko et al. 2012). (Note that we use the anatomical IFG-opercular parcel to constrain the definition of individual articulation-responsive fROIs instead of the parcel discovered in the GSS, because the former is larger and thus more likely to capture the relevant sets of voxels across individuals given the high variability in this part of the cortex.)

Below, we discuss the answers to our 3 research questions. (The data from this study are available at: <https://evlab.mit.edu/papers/artic>.)

1. Are Articulation-Responsive Regions Sensitive to Articulatory Complexity?

All articulation-responsive brain regions responded robustly to each of the 3 articulation conditions relative to the low-level fixation baseline (hard articulation: $ts > 5.7$, $Ps < 0.001$; easy articulation: $ts > 4.5$, $Ps < 0.001$; vowels: $ts > 3.6$, $Ps < 0.001$; Table 1 and Fig. 3). To evaluate the sensitivity of these regions to articulatory complexity, we asked—for each region—whether it showed (1) a reliably stronger response to the *hard* articulation than the *easy*

articulation condition, and (2) a reliably stronger response to the *easy* articulation than the *vowels* condition (which has minimal articulatory requirements).

All regions, except for the R IFGop fROI, showed robust effects for the *hard* articulation > *easy* articulation contrast ($ts > 2.9$, $Ps < 0.005$, with all but 3 of these fROIs surviving the stringent Bonferroni correction). Further, in all but 2 fROIs (R iPrCG, R IFGop), the *easy* articulation condition elicited a numerically higher response than the *vowel* condition, with 2 fROIs (L STG and R aSTG) showing this effect reliably. Thus, all articulation-responsive regions respond strongly during articulation, with conditions requiring more complex articulatory movements generally eliciting stronger responses, an expected functional signature of the articulation system (e.g., Bohland and Guenther 2006).

2. To What Extent are Articulation-Responsive Regions Functionally Selective for Speech Production?

To evaluate the degree of functional selectivity of the articulation-responsive brain regions, we asked—for each region—whether it showed a reliably stronger response to each of the 3 articulation conditions (*hard* articulation, *easy* articulation, and *vowels*) than to the nonarticulation conditions that share some features with speech production (*respiration* and *nonspeech oral-motor movements*). Both nonspeech conditions elicited reliably above-baseline responses in most of the regions except for some superior temporal fROIs (Bonferroni-surviving significant $ts > 3.7$, $Ps < 0.001$ for respiration, and $ts > 4.0$; $Ps < 0.001$ for nonspeech oral-motor movements; Table 2). However, the respiration and the nonspeech oral-motor movements conditions differed in that most fROIs showed a stronger response to the articulation conditions than the respiration condition (e.g., Bonferroni-surviving significant $ts > 4.6$, $Ps < 0.001$ for the *hard* articulation > *respiration* effect), but only some of the fROIs—namely, those in the superior temporal cortex (the L STG, L pSTG2, R aSTG, R pSTG, and R sPrCG fROIs)—showed a stronger response to *articulation* than to *nonspeech oral-motor movements*, with regions in the precentral gyrus showing the opposite pattern (Table 3 and Fig. 3). These results suggest that most articulation-responsive regions are not specific to speech production and are engaged to a similar extent by the production of nonspeech oral-motor movements.

Table 1 Responses of the articulation-responsive fROIs to (1) the *hard* articulation > *baseline* contrast; (2) the *easy* articulation > *baseline* contrast; (3) the *vowel* > *baseline* contrast; (4) the *hard* articulation > *easy* articulation contrast; and (5) the *easy* articulation > *vowels* contrast

fROI	Hard articulation > baseline	Easy articulation > baseline	Vowels > baseline	Hard > easy articulation	Easy articulation > vowels
L STG	$t = 10.33, P < 0.001^*$	$t = 9.95, P < 0.001^*$	$t = 6.98, P < 0.001^*$	$t = 5.57, P < 0.001^*$	$t = 4.67, P < 0.001^*$
L pSTG1	$t = 6.21, P < 0.001^*$	$t = 6.04, P < 0.001^*$	$t = 3.56, P < 0.005^*$	$t = 2.93, P < 0.005$	$t = 3.29, P < 0.005$
L pSTG2	$t = 5.65, P < 0.001^*$	$t = 4.53, P < 0.001^*$	$t = 4.77, P < 0.001^*$	$t = 5.67, P < 0.001^*$	$t = 1.17, n.s.$
R aSTG	$t = 10.05, P < 0.001^*$	$t = 8.53, P < 0.001^*$	$t = 6.33, P < 0.001^*$	$t = 5.71, P < 0.001^*$	$t = 4.08, P < 0.001^*$
R pSTG	$t = 5.69, P < 0.001^*$	$t = 5.54, P < 0.001^*$	$t = 6.06, P < 0.001^*$	$t = 3.07, P < 0.005$	$t = 1.30, n.s.$
L PrCG	$t = 9.16, P < 0.001^*$	$t = 9.04, P < 0.001^*$	$t = 8.03, P < 0.001^*$	$t = 3.84, P < 0.001^*$	$t = 0.84, n.s.$
R iPrCG	$t = 6.43, P < 0.001^*$	$t = 6.26, P < 0.001^*$	$t = 5.24, P < 0.001^*$	$t = 3.81, P < 0.001^*$	$t = -0.13, n.s.$
R sPrCG	$t = 6.99, P < 0.001^*$	$t = 7.18, P < 0.001^*$	$t = 5.76, P < 0.001^*$	$t = 3.25, P < 0.005$	$t = 1.14, n.s.$
SMA	$t = 8.15, P < 0.001^*$	$t = 8.36, P < 0.001^*$	$t = 6.14, P < 0.001^*$	$t = 3.90, P < 0.001^*$	$t = 0.75, n.s.$
L Cereb	$t = 7.91, P < 0.001^*$	$t = 7.29, P < 0.001^*$	$t = 4.67, P < 0.001^*$	$t = 3.73, P < 0.001^*$	$t = 0.78, n.s.$
R Cereb	$t = 9.30, P < 0.001^*$	$t = 9.00, P < 0.001^*$	$t = 5.62, P < 0.001^*$	$t = 4.02, P < 0.001^*$	$t = 1.32, n.s.$
L IFGop	$t = 8, P < 0.001^*$	$t = 7.25, P < 0.001^*$	$t = 7.06, P < 0.001^*$	$t = 3.78, P < 0.001^*$	$t = 0.73, n.s.$
R IFGop	$t = 6.18, P < 0.001^*$	$t = 6.31, P < 0.001^*$	$t = 5.98, P < 0.001^*$	$t = 2.06, n.s.$	$t = -1.32, n.s.$

All contrasts are estimated in data not used for defining the fROIs, as described in Methods. Degrees of freedom are 19. Asterisks denote effects that survive the Bonferroni correction for the number of fROIs (at $0.05/13 = 0.0038$).

In summary, both nonspeech conditions produced reliable above-baseline responses in most articulation-responsive brain regions. Furthermore, the nonspeech oral-motor movements condition produced a response that was as strong as, or stronger than, the articulation conditions in all the fROIs, except for the superior temporal ones and the L IFGop fROI. Thus, with the exception of the superior temporal fROIs and the fROI in the opercular portion of the left IFG, the hypotheses about the articulation-responsive regions' possible computations cannot be restricted to the articulatory domain.

3. What is the Relationship Between Articulation-Responsive Regions and the Domain-General Multiple Demand System?

First, the spatial WM task produced the expected behavioral and neural patterns. Behaviorally, participants were reliably faster and more accurate in the easy compared with the hard condition (RT: $t(19) = 9.46$, $P < 0.001$; accuracy: $t(19) = 12.94$, $P < 0.001$). Replicating prior work (e.g., Fedorenko et al. 2013), all

Table 2 Responses of the articulation-responsive fROIs to each of the nonspeech conditions (*respiration, nonspeech oral-motor (NSO) movements*) relative to the fixation baseline

fROIs	Resp > baseline	NSO > baseline
L STG	$t = 2.75$, $P < 0.01$	$t = 2.51$, $P < 0.05$
L pSTG1	$t = 4.09$, $P < 0.001^*$	$t = 4.01$, $P < 0.001^*$
L pSTG2	$t = 1.21$, n.s.	$t = 2.39$, $P < 0.05$
R aSTG	$t = 1.43$, n.s.	$t = 3.00$, $P < 0.005$
R pSTG	$t = 3.18$, $P < 0.005$	$t = 4.90$, $P < 0.001^*$
L PrCG	$t = 6.15$, $P < 0.001^*$	$t = 7.73$, $P < 0.001^*$
R iPrCG	$t = 4.39$, $P < 0.001^*$	$t = 8.02$, $P < 0.001^*$
R sPrCG	$t = 3.79$, $P < 0.001^*$	$t = 5.83$, $P < 0.001^*$
SMA	$t = 5.79$, $P < 0.001^*$	$t = 5.93$, $P < 0.001^*$
L Cereb	$t = 3.93$, $P < 0.001^*$	$t = 6.14$, $P < 0.001^*$
R Cereb	$t = 4.67$, $P < 0.001^*$	$t = 8.56$, $P < 0.001^*$
L IFGop	$t = 5.49$, $P < 0.001^*$	$t = 6.04$, $P < 0.001^*$
R IFGop	$t = 2.92$, $P < 0.005$	$t = 6.15$, $P < 0.001^*$

All contrasts are estimated in data not used for defining the fROIs, as described in Methods. Degrees of freedom are 19. Asterisks denote effects that survive the Bonferroni correction for the number of fROIs (at $0.05/13 = 0.0038$).

Table 3 Responses of the articulation-responsive fROIs to each of the articulation conditions (*hard articulation, easy articulation, and vowels*) relative to each of the 2 nonarticulation conditions (*respiration, nonspeech oral-motor (NSO) movements*)

fROIs	Hard artic > respiration	Easy artic > respiration	Vowel > respiration	Hard artic > NSO	Easy artic > NSO	Vowel > NSO
L STG	$t = 8.90$, $P < 0.001^*$	$t = 8.28$, $P < 0.001^*$	$t = 6.86$, $P < 0.001^*$	$t = 9.71$, $P < 0.001^*$	$t = 9.24$, $P < 0.001^*$	$t = 5.95$, $P < 0.001^*$
L pSTG1	$t = 5.27$, $P < 0.001^*$	$t = 2.05$, n.s.	$t = -0.74$, n.s.	$t = 3.67$, $P < 0.005$	$t = 2.41$, $P < 0.05$	$t = 0.13$, n.s.
L pSTG2	$t = 6.27$, $P < 0.001^*$	$t = 4.96$, $P < 0.001^*$	$t = 6.42$, $P < 0.001^*$	$t = 5.22$, $P < 0.001^*$	$t = 3.79$, $P < 0.001^*$	$t = 4.53$, $P < 0.001^*$
R aSTG	$t = 10.5$, $P < 0.001^*$	$t = 7.84$, $P < 0.001^*$	$t = 7.72$, $P < 0.001^*$	$t = 7.88$, $P < 0.001^*$	$t = 6.08$, $P < 0.001^*$	$t = 4.35$, $P < 0.001^*$
R pSTG	$t = 5.59$, $P < 0.001^*$	$t = 5.42$, $P < 0.001^*$	$t = 6.06$, $P < 0.001^*$	$t = 3.43$, $P < 0.005$	$t = 2.84$, $P < 0.05$	$t = 2.83$, $P < 0.05$
L PrCG	$t = 4.68$, $P < 0.001^*$	$t = 3.20$, $P < 0.005$	$t = 3.49$, $P < 0.005$	$t = -0.24$, n.s.	$t = -1.63$, n.s.	$t = -2.04$, n.s.
R iPrCG	$t = 2.87$, $P < 0.01$	$t = 1.78$, n.s.	$t = 3.32$, $P < 0.005$	$t = -2.06$, n.s.	$t = -3.39$, n.s.	$t = -3.68$, n.s.
R sPrCG	$t = 5.99$, $P < 0.001^*$	$t = 5.25$, $P < 0.001^*$	$t = 4.47$, $P < 0.001^*$	$t = -0.83$, n.s.	$t = -1.88$, n.s.	$t = -2.16$, n.s.
SMA	$t = 1.77$, n.s.	$t = 0.14$, n.s.	$t = -0.79$, n.s.	$t = 2.28$, $P < 0.05$	$t = 0.83$, n.s.	$t = 0.10$, n.s.
L Cereb	$t = 2.34$, $P < 0.05$	$t = 0.99$, n.s.	$t = 0.46$, n.s.	$t = 0.46$, n.s.	$t = -0.57$, n.s.	$t = -1.40$, n.s.
R Cereb	$t = 4.89$, $P < 0.001^*$	$t = 2.12$, $P < 0.05$	$t = 1.08$, n.s.	$t = 1.86$, n.s.	$t = -0.46$, n.s.	$t = -1.79$, n.s.
L IFGop	$t = 2.73$, $P < 0.05$	$t = 1.56$, n.s.	$t = 1.51$, n.s.	$t = 2.21$, $P < 0.05$	$t = 0.99$, n.s.	$t = 0.51$, n.s.
R IFGop	$t = 1.95$, n.s.	$t = 1.40$, n.s.	$t = 4.25$, $P < 0.001^*$	$t = 1.39$, n.s.	$t = 2.43$, $P < 0.05$	$t = 1.43$, n.s.

All contrasts are estimated in data not used for defining the fROIs, as described in Methods. Degrees of freedom are 19. Asterisks denote effects that survive the Bonferroni correction for the number of fROIs (at $0.05/13 = 0.0038$).

MD fROIs showed stronger responses during the hard than the easy condition, in data not used for fROI definition as discussed in Methods ($ts > 4.3$, $Ps < 0.0001$; see also Fig. 4).

To evaluate the relationship between articulation-responsive fROIs and the MD system, we performed several analyses. First, we asked whether the articulation-responsive regions showed a reliably above-baseline response to the hard and easy spatial WM conditions. Only the SMA, L and R Cereb, and L and R IFGop fROIs showed reliable above-baseline responses to the *hard spatial WM* condition ($ts > 3.6$, $Ps < 0.001$), with a subset further showing reliable responses to the *easy spatial WM* condition ($ts > 3.8$, $Ps < 0.001$; Table 4 and Fig. 3). In each of these 5 regions, the *hard spatial WM* condition elicited a stronger response than the *easy spatial WM* condition, reliably so in 2 of the fROIs (L Cereb, R IFGop; $ts > 3.8$, $Ps < 0.001$). None of the other articulation-responsive fROIs responded during the spatial WM task; nor were they sensitive to the task difficulty manipulation, implying some selectivity relative to general demanding tasks.

Second, we directly compared the average activation for the 2 articulation conditions (*hard articulation, easy articulation*) to the average activation during the spatial WM conditions (*hard spatial WM, easy spatial WM*) in each articulation-responsive fROI. All regions, except for the R IFGop fROI, responded reliably more strongly to the articulation conditions than the spatial WM conditions ($ts > 4.7$, $Ps < 0.001$; Table 4 and Fig. 3), further supporting selectivity relative to executive tasks, and thus separability from the nearby MD network (Duncan 2010, 2013).

Third, in a complementary analysis, we examined the MD fROIs' (defined by the spatial WM task) responses to articulation. A number of the MD regions (in particular, L and R PrCG fROIs, L and R SMA fROIs, and L ParInf fROI) responded reliably above baseline to each of the articulation conditions (Bonferroni-surviving significant ts : *hard articulation*: $ts > 3.7$, $Ps < 0.001$; *easy articulation*: $ts > 3.6$, $Ps < 0.001$; Table 5 and Fig. 4), suggesting that these domain-general brain regions may contribute in some (albeit nonspeech-specific) way to speech production. However, only the R Ins MD fROI showed a reliable *hard > easy articulation* effect. Thus, the articulation-responsive MD fROIs do not show the functional signature of articulation-responsive fROIs, that is, robust sensitivity to articulatory complexity (Table 5).

Finally, given that some macroanatomical regions appear to exhibit some sensitivity to both articulation and general cognitive effort (estimated here with a spatial WM task), we

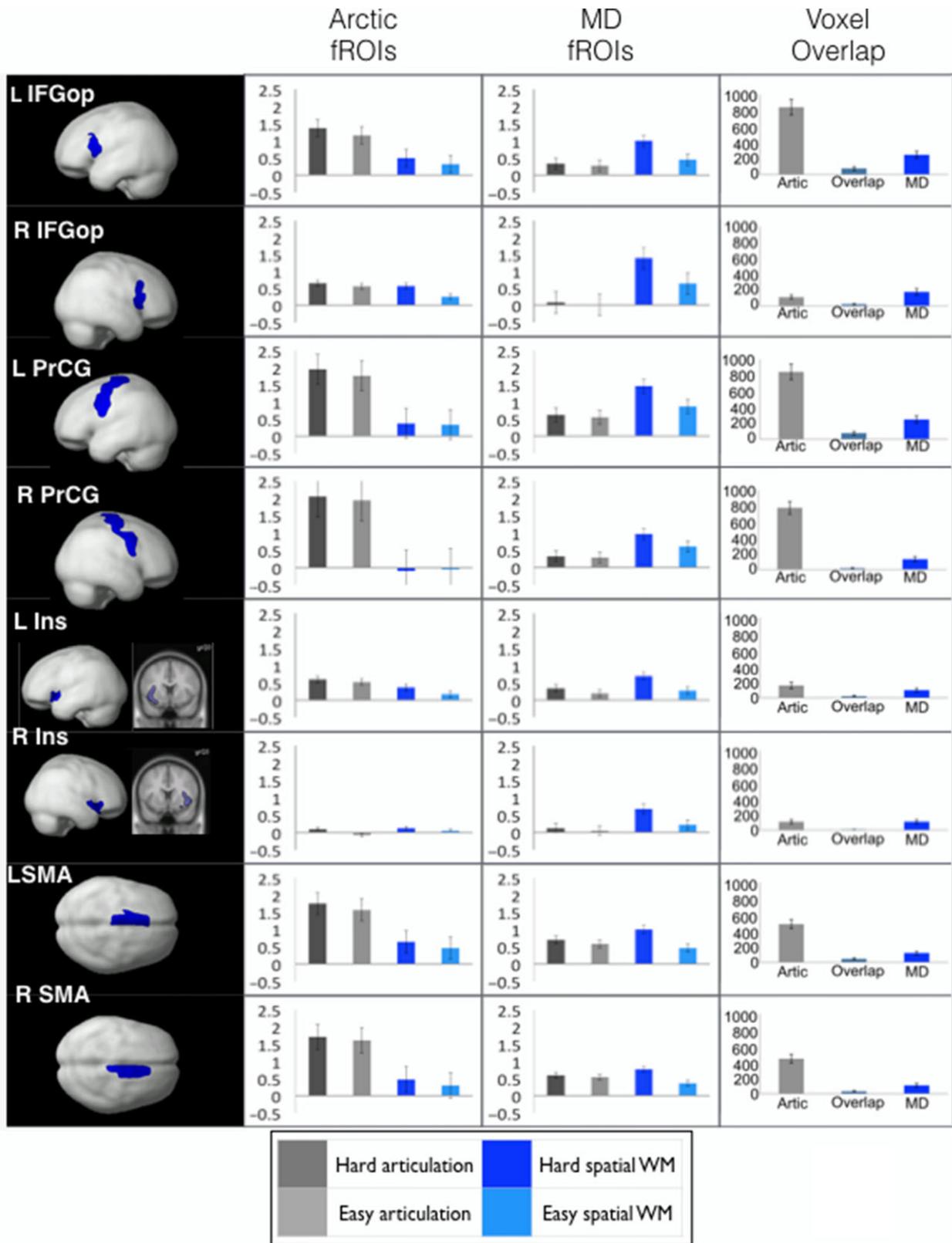


Figure 4. Responses to the articulation and spatial WM conditions as a function of fROI type within the macroanatomical regions that show sensitivity to both articulation and general cognitive demand. *Left column:* The anatomical regions (a subset of the MD parcels), which include L and R IFGop, PrCG, Insula, and SMA. *Middle columns:* Responses to the articulation and spatial WM conditions in the articulation-responsive fROIs (defined by the hard articulation > fixation contrast; the “Arctic fROIs” column) and in the MD fROIs (defined by the hard > easy spatial WM contrast; the “MD fROIs” column). *Right column:* Voxel counts for the hard articulation > fixation contrast, the hard > easy spatial WM contrast at the $P < 0.001$, uncorrected at the whole-brain level threshold, and voxels that show significant effects for both contrasts.

Table 4 Responses of the articulation-responsive fROIs (1) to each of the spatial WM conditions relative to the fixation baseline; (2) to the *hard* > *easy spatial* WM contrast; and (3) to the articulation conditions (*hard* articulation, *easy* articulation) relative to the spatial WM conditions (*hard*, *easy*)

fROI	Hard spatial WM > baseline	Easy spatial WM > baseline	Hard spatial WM > easy spatial WM	Articulation > spatial WM
L STG	t = -2.62, n.s.	t = -1.53, n.s.	t = -2.27, n.s.	t = 10.52, P < 0.001*
L pSTG1	t = -0.48, n.s.	t = 1.21, n.s.	t = -2.50, n.s.	t = 6.13, P < 0.001*
L pSTG2	t = -3.37, n.s.	t = -2.59, n.s.	t = -2.25, n.s.	t = 5.25, P < 0.001*
R aSTG	t = -2.69, n.s.	t = -1.49, n.s.	t = -2.68, n.s.	t = 9.07, P < 0.001*
R pSTG	t = -1.48, n.s.	t = -0.56, n.s.	t = -1.98, n.s.	t = 5.57, P < 0.001*
L PrCG	t = -0.50, n.s.	t = 0.76, n.s.	t = -3.48, n.s.	t = 8.42, P < 0.001*
R iPrCG	t = -3.48, n.s.	t = -1.69, n.s.	t = -3.92, n.s.	t = 6.89, P < 0.001*
R sPrCG	t = -1.96, n.s.	t = -1.21, n.s.	t = -2.24, n.s.	t = 8.06, P < 0.001*
SMA	t = 3.64, P < 0.001*	t = 3.43, P < 0.005	t = 2.33, P < 0.05	t = 6.25, P < 0.001*
L Cereb	t = 3.97, P < 0.001*	t = 2.88, P < 0.005	t = 4.18, P < 0.001*	t = 4.78, P < 0.001*
R Cereb	t = 4.34, P < 0.001*	t = 3.89, P < 0.001*	t = 3.56, P < 0.005	t = 5.79, P < 0.001*
L IFGop	t = 3.78, P < 0.001*	t = 3.89, P < 0.001*	t = 2.52, P < 0.05	t = 5.62, P < 0.001*
R IFGop	t = 3.71, P < 0.001*	t = 2.63, P < 0.01	t = 3.85, P < 0.001*	t = 1.73, n.s.

All contrasts are estimated in data not used for defining the fROIs, as described in Methods. Degrees of freedom are 19. Asterisks denote effects that survive the Bonferroni correction for the number of fROIs (at 0.05/13 = 0.0038).

Table 5 Responses of the MD fROIs (1) to each articulation condition (*hard*, *easy*) relative to the fixation baseline; and (2) to the *hard* > *easy* articulation contrast

MD fROI	Hard articulation > baseline	Easy articulation > baseline	Hard Artic. > easy Artic.
L IFGop	t = 2.28, P < 0.05	t = 2.18, P < 0.05	t = 0.87, P = n.s.
R IFGop	t = 0.72, P = n.s.	t = 0.12, P = n.s.	t = 1.06, P = n.s.
L MFG	t = 2.68, P < 0.01	t = 2.49, P < 0.05	t = 0.89, P = n.s.
R MFG	t = 2.39, P < 0.01	t = 1.41, P = n.s.	t = 1.86, P < 0.05
L MFGorb	t = -0.57, P = n.s.	t = -1.40, P = n.s.	t = 1.31, P = n.s.
R MFGorb	t = 0.30, P = n.s.	t = -1.08, P = n.s.	t = 2.04, P < 0.05
L PrCG	t = 4.41, P < 0.001*	t = 4.24, P < 0.001*	t = 1.26, P = n.s.
R PrCG	t = 4.24, P < 0.001*	t = 3.69, P < 0.001*	t = 0.96, P = n.s.
L Ins	t = 3.22, P < 0.005	t = 1.73, P < 0.05	t = 2.70, P < 0.01
R Ins	t = 1.27, P = n.s.	t = -1.00, P = n.s.	t = 3.66, P < 0.001*
L SMA	t = 6.96, P < 0.001*	t = 5.69, P < 0.001*	t = 1.81, P < 0.05
R SMA	t = 4.43, P < 0.001*	t = 4.01, P < 0.001*	t = 1.01, P = n.s.
L ParInf	t = 3.70, P < 0.001*	t = 3.81, P < 0.001*	t = 0.72, P = n.s.
R ParInfo	t = 1.67, P = n.s.	t = 1.16, P = n.s.	t = 0.80, P = n.s.
L ParSup	t = 0.74, P = n.s.	t = 0.16, P = n.s.	t = 1.20, P = n.s.
R ParSup	t = -2.23, P = n.s.	t = -2.10, P = n.s.	t = 1.29, P = n.s.
L ACC	t = -1.31, P = n.s.	t = -1.83, P = n.s.	t = 0.84, P = n.s.
R ACC	t = -0.87, P = n.s.	t = -1.30, P = n.s.	t = 1.88, P < 0.05

fROI names: IFGop, inferior frontal gyrus, opercular portion; MFG, middle frontal gyrus; MFGorb, middle frontal gyrus, orbital portion; PrCG, precentral gyrus; SMA, supplementary motor area; InfPar, inferior parietal cortex; SupPar, superior parietal cortex; ACC, anterior cingulate (see e.g., Fedorenko et al. 2013, for more details). All contrasts are estimated in data not used for defining the fROIs, as described in Methods. Degrees of freedom are 19. Asterisks denote effects that survive the Bonferroni correction for the number of fROIs (at 0.05/18 = 0.0027).

directly examined the relationship between articulation-responsive fROIs and domain-general MD system fROIs within those broad areas (using the masks from the Tzourio-Mazoyer et al. 2002 atlas to constrain the definition of individual fROIs). Based on the articulation-responsive regions discovered by the GSS analysis in this study, we included L and R PrecG and L and R SMA. Based on prior studies of articulation (e.g., Dronkers 1996; Wise et al. 1999; Bohland and Guenther 2006), we additionally examined L and R IFGop and L and R Insula. Within each of those regions, we defined 2 fROIs: an articulation-responsive fROI (using the *hard* articulation > fixation contrast) and an MD fROI (using the *hard* > *easy spatial* WM contrast) in each participant, and then examined the responses of these fROIs to the articulation and spatial WM conditions (using data

not used for fROI definition, as described in Methods). For each region, we performed a 2 (fROI type: articulation, MD) × 2 (Task: Articulation, Spatial WM) × 2 (Difficulty: *Hard*, *Easy*) analysis of variance. An interaction between fROI type and Task would indicate distinct functional profiles for the articulation fROIs versus MD fROIs within these broad anatomical areas. We further examined the overlap between voxels that show a reliable *hard* articulation > baseline effect and voxels that show a reliable *hard* spatial WM > *easy* spatial WM effect (at the fixed threshold of P < 0.001, uncorrected whole-brain level).

We observed a highly reliable interaction between fROI type and Task in each region ($F_s > 6.10$, $P_s < 0.05$; Fig. 4 and Table 6). The articulation-responsive fROIs showed reliably stronger responses to the 2 articulation conditions than the spatial WM

Table 6 Results from the fROI type \times Task and fROI type \times Task \times Difficulty interactions for each ROI

ROI	F (1,19)	P
L IFGop		
fROI Type \times Task	62.08	<0.001
fROI Type \times Task \times Difficulty	39.31	<0.001
R IFGop		
fROI Type \times Task	65.24	<0.001
fROI Type \times Task \times Difficulty	28.60	<0.001
L PrCG		
fROI Type \times Task	170.52	<0.001
fROI Type \times Task \times Difficulty	58.11	<0.001
R PrCG		
fROI Type \times Task	245.66	<0.001
fROI Type \times Task \times Difficulty	51.65	<0.001
L Ins		
fROI Type \times Task	14.87	0.001
fROI Type \times Task \times Difficulty	2.40	n.s.
R Ins		
fROI Type \times Task	6.10	<0.05
fROI Type \times Task \times Difficulty	20.92	<0.001
L SMA		
fROI Type \times Task	59.43	<0.001
fROI Type \times Task \times Difficulty	28.61	<0.001
R SMA		
fROI Type \times Task	50.20	<0.001
fROI Type \times Task \times Difficulty	11.60	<0.005

conditions, and the MD fROIs showed reliably stronger responses to the spatial WM conditions than the articulation conditions. Furthermore, a 3-way interaction among fROI type, Task, and Difficulty was obtained in all regions except for the L Ins ($F_s > 11.60$, $P_s < 0.005$), suggesting that the difficulty effect is larger for the articulation conditions in the articulation-responsive fROIs, and for the spatial WM conditions in the MD fROIs. Finally, the voxel overlap analysis revealed that articulation-responsive voxels (significant for the *hard articulation > fixation* contrast at $P < 0.001$ uncorrected whole-brain level) were mostly nonoverlapping with the MD voxels (significant for the *hard > easy spatial WM* contrast at $P < 0.001$ uncorrected whole-brain level). Voxels that showed sensitivity to both articulation and general cognitive demand constituted between 1.8% and 9.6% of the total voxels activated by either the *hard articulation > fixation* contrast or the *hard > easy spatial WM* contrast. (Note that the greater number of articulation-responsive voxels than MD voxels is simply due to the use of a functionally broader contrast, with the fixation baseline for the former.)

In tandem, these analyses suggest that although parts of the MD system exhibit some (relatively weak) response during articulation, in line with some role for domain-general executive resources in speech production, there also exist articulation-responsive regions—in close proximity to those MD regions—that show little or no response to nonspeech demanding executive tasks (Fig. 4).

Discussion

Speech production requires an orchestrated series of cognitive and motor events: Once a message is conceived and phonological planning has begun, fluent speech requires coordination of respiration, phonation, articulation, resonance, and prosody (e.g., Darley 1969). Using an individual-subjects functional localization approach (e.g., Fedorenko et al. 2010), we first identified a

set of brain regions that respond during an articulation task (pseudoword repetition). These regions corresponded well with a set of regions previously implicated in speech production based on patient (e.g., Hillis et al. 2004; Richardson et al. 2012; Graff-Radford et al. 2014; Hickok et al. 2014; Basilakos et al. 2015; Itabashi et al. 2016) and brain imaging studies (e.g., Wise et al. 1999; Bohland and Guenther 2006; Eickhoff et al. 2009; Fedorenko et al. 2015) and included portions of the precentral gyrus, SMA, inferior frontal cortex, superior temporal cortex, and cerebellum. All of these regions showed robust sensitivity to articulatory complexity (see Table 7, below, for a summary).

We then evaluated the degree to which articulation-responsive brain regions—defined functionally in each individual participant—are selective for speech production. In Experiment 2, we examined these regions' responses to non-articulation tasks that share features with speech production, and found a clear functional dissociation between the superior temporal fROIs and the left IFGop fROI, which showed selective responses to articulation, and the fROIs in the precentral gyrus, the SMA, and the cerebellum, which responded as strongly during the production of nonspeech oral-motor movements as during articulation, and showed a substantial response during the respiration condition. Finally, in Experiment 3, we examined the relationship between the articulation-responsive regions and the domain-general MD regions (e.g., Duncan 2010, 2013) and found that the articulation regions were spatially and functionally distinct from the nearby MD regions, in spite of the fact that articulation is a complex cognitively demanding goal-directed behavior (see Table 7). In the remainder of the paper, we discuss how the current results inform the architecture of the human articulation system.

Ubiquitous Sensitivity of Articulation-Responsive Regions to Articulatory Complexity

Fluent speech requires fast sequencing of motor movements of the articulatory apparatus. Articulatory complexity can be conceptualized in different ways, from the length of the sequence, to its familiarity, to the particular sounds involved, and the speed at which the sequence has to be produced. Prior studies have explored and documented the effects of these different factors on the observed behavior and brain activity (e.g., Price 2012), although in most prior neuroimaging studies (e.g., Segawa et al. 2015), observations of more neural activity for the more complex condition(s) are difficult to interpret as reflecting specific sensitivity of the articulation regions versus nearby “domain-general” effort-sensitive MD regions without identifying those domain-general regions in the same participants functionally; see Fedorenko et al. 2013 for discussion). What seems to cause the greatest difficulty behaviorally, both in development and in disordered speech production, are sequences that involve clusters of consonant sounds, as their production requires quick transitions between articulator positions (e.g., Romani et al. 2002; Romani and Galluzzi 2005; Staiger and Ziegler 2008). We therefore manipulated complexity in this way.

All articulation-responsive regions were robustly sensitive to articulatory complexity, with stronger responses to conditions that require more complex articulatory movements. Producing CCVCC-CV pseudowords elicited greater responses than simple CV-CV pseudowords, which in turn produced a greater response than vowels. Thus, producing harder-to-articulate sequences requires more neural activity. (Interestingly, in spite of the observed differences in the mean response for the hard and easy articulation conditions, the *fine-grained activation patterns*

Table 7 A summary of the key results for the articulation fROIs

fROI	Sensitivity to articulatory complexity	Selectivity relative to respiration	Selectivity relative to NSO movements	Selectivity relative to a non-speech demanding task	Sensitivity to general cognitive effort
L STG	✓*	✓*	✓*	✓*	X
L pSTG1	✓	✓*	✓	✓*	X
L pSTG2	✓*	✓*	✓*	✓*	X
R aSTG	✓*	✓*	✓*	✓*	X
R pSTG	✓	✓*	✓	✓*	X
L PrCG	✓*	✓*	X	✓*	X
R iPrCG	✓*	✓	X	✓*	X
R sPrCG	✓	✓*	X	✓*	X
SMA	✓*	X	✓	✓*	✓
L Cereb	✓*	✓	X	✓*	✓*
R Cereb	✓*	✓*	X	✓*	✓
L IFGop	✓*	✓	✓	✓*	✓
R IFGop	X	X	X	X	✓*

Sensitivity to articulatory complexity is assessed by the *hard articulation* > *easy articulation* contrast (Table 1). Selectivity relative to respiration and NSO movements are assessed by the *hard articulation* > *respiration* and *hard articulation* > *NSO* contrasts, respectively (Table 3). Selectivity relative to a nonspeech demanding task is assessed by the *articulation* > *spatial WM* contrast (Table 4). Finally, sensitivity to general cognitive effort is assessed by the *hard spatial WM* > *easy spatial WM* contrast (Table 4). A checkmark with an asterisk (darker gray) indicates a significant effect that survived the Bonferroni correction; a checkmark without an asterisk (lighter gray) indicates an effect at $P < 0.05$ or lower that did not survive the Bonferroni correction; a cross (white) indicates a nonsignificant effect.

were similar between these conditions across the articulation network, as shown in the Supplementary Information, suggesting that these brain regions are engaged in a similar way by these conditions.)

Differences Among Articulation-Responsive Regions with Respect to Their Selectivity for Speech Production

Prior studies have reported activity in premotor and motor cortices during speech production (e.g., Murphy et al. 1997; Wise et al. 1999; Lotze et al. 2000; Brown et al. 2005; Bohland and Guenther 2006; Guenther et al. 2006; Sörös et al. 2006), but also during respiration (e.g., Mckay et al. 2003; Loucks et al. 2007) and diverse nonspeech oral-motor behaviors including swallowing (e.g., Humbert and Robbins 2007), whistling (e.g., Dresel et al. 2005), and articulator movements (e.g., Grabski et al. 2012). One fMRI study directly compared activation for speech production and the production of nonspeech vocal tract gestures and reported overlap (Chang et al. 2009). However, the comparisons were performed at the group level, which can overestimate activation overlap (e.g., Nieto-Castañón and Fedorenko 2012). We therefore defined articulation-responsive regions in each brain individually and asked whether these exact sets of voxels respond during respiration and nonspeech oral-motor movement production. Most articulation-responsive fROIs responded at least as strongly during the production of nonspeech oral-motor movements as during articulation, and quite strongly during the respiration condition.

A similarly strong response to articulation and nonspeech oral-motor movement production in the articulation-responsive regions within the precentral gyrus is consistent with prior findings of somatotopic organization of sensorimotor cortices (e.g., Penfield and Rasmussen 1950; Lotze et al. 2000; Hesselmann et al. 2004; Pulvermüller et al. 2006; Grabski et al. 2012).

Furthermore, a recent electrocorticography (ECoG) study showed that ventral sensorimotor cortex is organized by key articulator features and articulator positions (Bouchard et al. 2013; see also Guenther 2016, for a proposal that the PrecG contains the articulator map). Plausibly, the same circuits that implement the type of articulator and its position would be engaged whenever that articulator is required, whether for speech or nonspeech movements, in line with proposals that speech motor and nonspeech oral-motor control processes form an integrated system of motor control (e.g., Ballard et al. 2003, 2009; cf. Ziegler 2003a, 2003b, 2006). (Interestingly, despite similar mean responses, fine-grained activation patterns were highly distinct between the articulation conditions and the nonspeech oral-motor movement condition, as revealed by multivariate analyses, shown in the SI, suggesting that these conditions engage the same regions, but in different ways.)

An architecture where the same brain regions support control of articulators in the context of both speech and nonspeech movements is consistent with evidence that speech production deficits are often accompanied by general difficulties with oral-motor control (e.g., Ballard et al. 2000; Robin et al. 2008). However, cases of speech difficulties without general oral-motor problems have also been reported (e.g., Kwon et al. 2013; Whiteside et al. 2015). Might there be any parts of the speech articulation network selectively engaged in speech production relative to nonspeech movements?

Articulation-responsive regions in the cerebellum, like the PrCG fROIs, responded strongly during nonspeech oral-motor movement production. In contrast, most articulation regions in the superior temporal cortex exhibited selectivity for speech, with the articulation conditions eliciting stronger responses than the nonspeech conditions. This result suggests that these regions respond only when we produce, or prepare to produce, speech output. Current models of speech production

hypothesize that these superior temporal regions support phonological planning (e.g., Hickok 2009 for review) and/or monitor the outputs of the motor system as needed for motor control during vocal production (e.g., Guenther 2006; Rauschecker and Scott 2009). However, these regions are not likely candidates for eliciting selective speech articulation deficits, as evidenced by patient and intraoperative stimulation studies. Are there other, more plausible, candidates?

Although our whole-brain GSS analysis did not discover a robust articulation-responsive region in posterior IFG (plausibly because of especially high interindividual variability in this part of the brain; e.g., Fischl et al. 2008; Frost and Goebel 2012), the analysis where we used an anatomical parcel for opercular IFG and searched for articulation-responsive voxels within it—in each participant individually—revealed highly robust and replicable responses to articulation, as well as sensitivity to articulatory complexity in the L IFGop fROI ($F(1,19) = 39.3$, $P < 0.001$; Figs 3 and 4). The L IFGop fROI further showed some degree of selectivity for speech production (Figs 3 and 4), with the hard articulation condition producing a stronger response than both respiration ($t(19) = 2.73$, $P < 0.05$) and nonspeech oral-motor movements ($t(19) = 2.21$, $P < 0.05$) (see Table 3). In contrast, in the R IFGop fROI, the nonspeech oral-motor movement condition produced the strongest response.

Regions in the vicinity of left posterior IFG have been associated with AOS (e.g., Hillis et al. 2004; Richardson et al. 2012), and most current models link this region with speech-specific functions, such as implementing motor syllable programs (Hickok 2014) or containing portions of the speech sound map (Guenther 2016), but the precise contribution of this region to articulation remains debated.

A recent ECoG study (Flinker et al. 2015) demonstrated that sites within posterior IFG appear to mediate information flowing from the temporal cortex (where phonological planning likely occurs; e.g., Hickok 2009) to motor regions. Given that during actual speech, only the motor regions were active, and sites in posterior IFG were silent, Flinker et al. hypothesized that the latter prepare an articulatory code that is sent to the motor cortex where it is implemented. Although this study provided important information about the division of labor between articulation-responsive regions in posterior IFG versus premotor/motor regions, it left open the question of whether sites in posterior IFG are selectively engaged during speech planning or whether they also support the planning of nonspeech oral-motor movements, thus leaving it unclear whether damage to this region could result in selective speech articulation deficits.

Another recent study used cortical cooling intraoperatively (Long et al. 2016) and found that interfering with neural activity in left posterior IFG affected the timing of speech, but not its quality (which was instead affected by cooling the speech motor cortex). The authors suggested that left posterior IFG may support sequence generation for speech production (see also Gelfand and Bookheimer 2003; Clerget et al. 2011; Uddén and Bahlmann 2012), although again, whether this region is selective for sequencing during articulation or also nonspeech motor behaviors was not established (e.g., Kimura 1982; Goodale 1988).

Thus, given some degree of speech selectivity exhibited by the articulation-responsive portion of the left IFGop in the current study, it appears possible that restricted damage to this region would lead to a selective deficit in speech production, but not to difficulties with nonspeech oral-motor movements. However, it remains critical to establish this

selectivity with methods that allow precise targeting of particular brain sites, ideally using functional localization, and afford causal inferences, such as electrical brain stimulation (e.g., Parvizi et al. 2012; cf. Borchers et al. 2012) or cooling (e.g., Long et al. 2016).

Another difference emerged between the articulation regions within the precentral gyrus versus posterior IFG: In the degree of lateralization. The PrCG articulation fROIs are robustly present in both hemispheres with no reliable differences in the strength of the response to the articulation conditions (Fig. 4, Table 5). Consistent with these robust bilateral activations, Tate et al. (2014; see also Cogan et al. 2014; cf. Long et al. 2016) found that intraoperative stimulation of both left and right precentral and postcentral gyri produced speech production deficits similar to those observed in dysarthria. In contrast, the IFGop articulation fROIs are highly asymmetrical with the response in the left hemisphere being ~3 times stronger than in the right hemisphere. This hemispheric asymmetry is in line with the fact that left, but not right, hemisphere damage to posterior IFG leads to articulation deficits (e.g., Damasio 1992).

The Relationship Between the Articulation System and the Domain-General multiple demand System

Given that MD regions lie in close proximity to some of the articulation regions, and are sensitive to effort across diverse tasks (e.g., Duncan and Owen 2000; Fedorenko et al. 2013; Hugdahl et al. 2015), we tested whether portions of the activation landscape for articulation may result from the engagement of highly domain-general processes, like attention or cognitive control. We confirmed that articulation and MD regions lie adjacent to each other within several macroanatomical regions; however, these sets of regions are almost entirely nonoverlapping at the level of individual participants. These results once again highlight the fact that functionally distinct subregions often lie in close proximity to one another within the same macroanatomical areas, and thus discussing activation peaks from across different manipulations/studies at the level of these broad areas can be misleading.

To conclude, we examined the brain regions of the articulation network and found robust and ubiquitous sensitivity to articulatory complexity, an expected marker of articulation-responsive regions. Further, articulation-responsive parts of the left posterior IFG and regions in the superior temporal cortex showed selectivity for articulation over nonspeech conditions, but bilateral regions in the precentral gyrus responded strongly during both articulation and the production of nonspeech oral-motor movements. Thus, the former regions may support speech-selective computations, but the latter are not likely to do so. Finally, despite close spatial proximity and despite articulation being a complex goal-directed behavior, articulation-responsive regions are functionally distinct from the domain-general regions of the fronto-parietal MD network whose regions are sensitive to diverse demanding tasks. Of course, these domain-general regions may still be important for some aspects of speech perception and/or production, but they likely support distinct computations from those implemented in the articulation regions. These different contributions need to be taken into account when further developing models of the human articulation system.

Supplementary Material

Supplementary data are available at *Cerebral Cortex* online.

Funding

National Institute on Deafness and Other Communication Disorders (grant DC009571 to J.F.), National Institute of Child Health and Human Development (grant HD057522 to E.F.); E.F. was also supported by a grant from the Simons Foundation to the Simons Center for the Social Brain at MIT, and by grants ANR-11-LABX-0036 (BLRI) and ANR-11-IDEX-0001-02 (A*MIDEX).

Notes

We thank Peter Graff and Morris Alper for help in the construction of pseudowords, Taylor Hanayik for help with recording the materials, Chris Rorden and Alfonso Nieto-Castañon for help with some preprocessing steps, Brianna Pritchett for help with the multivariate analyses, and Zach Mineroff for help with creating a website for this manuscript. *Conflict of Interest*: None declared.

References

- Amunts K, Lenzen M, Friederici AD, Schleicher A, Morosan P, Palomero-Gallagher N, Zilles K. 2010. Broca's region: novel organizational principles and multiple receptor mapping. *PLoS Biol.* 8:e1000489.
- Amunts K, Schleicher A, Bürgel U, Mohlberg H, Uylings H, Zilles K. 1999. Broca's region revisited: cytoarchitecture and intersubject variability. *J Comp Neurol.* 412:319–341.
- Baldo JV, Wilkins DP, Ogar J, Willock S, Dronkers N. 2011. Role of the precentral gyrus of the insula in complex articulation. *Cortex.* 47:800–807.
- Ballard K, Robin D, Folkins J. 2003. An integrative model of speech motor control: a response to Ziegler. *Aphasiology.* 17:37–48.
- Ballard KJ, Granier JP, Robin DA. 2000. Understanding the nature of apraxia of speech: theory, analysis and treatment. *Aphasiology.* 14:969–995.
- Ballard KJ, Solomon NP, Robin DA, Moon JB, Folkins JW, McNeil M. 2009. Nonspeech assessment of the speech production mechanism. New York: Thieme Medical.
- Basilakos A, Rorden C, Bonilha L, Moser D, Fridriksson J. 2015. Patterns of poststroke brain damage that predict speech production errors in apraxia of speech and aphasia dissociate. *Stroke.* 46:1561–1566.
- Blank I, Kanwisher N, Fedorenko E. 2014. A functional dissociation between language and multiple-demand systems revealed in patterns of BOLD signal fluctuations. *J Neurophysiol.* 112:1105–1118.
- Blumstein SE, Baum SR. 2016. Neurobiology of speech production: Perspective from neuropsychology and neurolinguistics. In: Hickok G, Small SL, editors. *Neurobiology of language*. Boston: Elsevier. p. 689–697.
- Bohland JW, Guenther FH. 2006. An fMRI investigation of syllable sequence production. *Neuroimage.* 32:821–841.
- Borchers S, Himmelfach M, Logothetis N, Karnath H-O. 2012. Direct electrical stimulation of human cortex—the gold standard for mapping brain functions? *Nat Rev Neurosci.* 13:63–70.
- Bouchard KE, Mesgarani N, Johnson K, Chang EF. 2013. Functional organization of human sensorimotor cortex for speech articulation. *Nature.* 495:327–332.
- Broca P. 1861. Remarques sur le siège de la faculté du langage articulé, suivies d'une observation d'aphémie (perte de la parole). [Translated in Eling (1994), p. 41–46]. *Bulletins de la Société Anatomique* 6.
- Browman CP, Goldstein L. 1995. Dynamics and articulatory phonology. In: Port R, van Gelder T, editors. *Mind as motion*. Cambridge (MA): MIT Press. p. 175–193.
- Brown S, Ingham RJ, Ingham JC, Laird AR, Fox PT. 2005. Stuttered and fluent speech production: an ALE meta-analysis of functional neuroimaging studies. *Hum Brain Mapp.* 25:105–117.
- Buchwald A. 2014. Phonetic processing. In: Goldrick MA, Ferreira V, Miozzo M, editors. *The Oxford handbook of language production*. New York: Oxford University Press. p. 245–258.
- Chang S-E, Kenney MK, Loucks TM, Poletto CJ, Ludlow CL. 2009. Common neural substrates support speech and non-speech vocal tract gestures. *Neuroimage.* 47:314–325.
- Clerget E, Badets A, Duqué J, Olivier E. 2011. Role of Broca's area in motor sequence programming: a cTBS study. *Neuroreport.* 22:965–969.
- Cogan GB, Thesen T, Carlson C, Doyle W, Devinsky O, Pesaran B. 2014. Sensory-motor transformations for speech occur bilaterally. *Nature.* 507:94–98.
- Damasio AR. 1992. Aphasia. *N Engl J Med.* 326:531–539.
- Darley F. 1969. The classification of output disturbances in neurogenic communication disorders. Paper Presented at the American Speech and Hearing Association Annual Conference.
- Darley FL, Aronson AE, Brown JR. 1975. *Motor speech disorders*. Philadelphia, PA: W.B. Saunders Co.
- Deen B, Koldewyn K, Kanwisher N, Saxe R. 2015. Functional organization of social perception and cognition in the superior temporal sulcus. *Cereb Cortex.* bhv111:1–14.
- Dell GS. 2014. Phonemes and production. *Lang Cogn Neurosci.* 29:30–32.
- Dresel C, Castrop F, Haslinger B, Wohlschlaeger AM, Hennenlotter A, Ceballos-Baumann AO. 2005. The functional neuroanatomy of coordinated orofacial movements: sparse sampling fMRI of whistling. *Neuroimage.* 28:588–597.
- Dronkers N. 1996. A new brain region for coordinating speech articulation. *Nature.* 384:159–161.
- Dronkers N, Ogar J. 2004. Brain areas involved in speech production. *Brain.* 127:1461–1462.
- Dronkers N, Ogar J, Willock S, Wilkins DP. 2004. Confirming the role of the insula in coordinating complex but not simple articulatory movements. *Brain Lang.* 91:23–24.
- Duffy J. 2005. *Motor speech disorders: substrates, differential diagnosis and management*. St. Louis, MO: Mosby.
- Duncan J. 2010. The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends Cogn Sci.* 14:172–179.
- Duncan J. 2013. The structure of cognition: attentional episodes in mind and brain. *Neuron.* 80:35–50.
- Duncan J, Owen AM. 2000. Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci.* 23:475–483.
- Eickhoff SB, Heim S, Zilles K, Amunts K. 2009. A systems perspective on the effective connectivity of overt speech production. *Phil Trans R Soc A.* 367:2399–2421.
- Fedorenko E, Duncan J, Kanwisher N. 2012. Language-selective and domain-general regions lie side by side within Broca's area. *Curr Biol.* 22:2059–2062.

- Fedorenko E, Duncan J, Kanwisher N. 2013. Broad domain generality in focal regions of frontal and parietal cortex. *Proc Natl Acad Sci USA*. 110:16616–16621.
- Fedorenko E, Fillmore P, Smith K, Bonilha L, Fridriksson J. 2015. The superior precentral gyrus of the insula does not appear to be functionally specialized for articulation. *J Neurophysiol*. 113:2376–2382.
- Fedorenko E, Hsieh P-J, Nieto-Castañón A, Whitfield-Gabrieli S, Kanwisher N. 2010. New method for fMRI investigations of language: defining ROIs functionally in individual subjects. *J Neurophysiol*. 104:1177–1194.
- Fedorenko E, Thompson-Schill SL. 2014. Reworking the language network. *Trends Cogn Sci*. 18:120–126.
- Fischl B, Rajendran N, Busa E, Augustinack J, Hinds O, Yeo BT, Mohlberg H, Amunts K, Zilles K. 2008. Cortical folding patterns and predicting cytoarchitecture. *Cereb Cortex*. 18:1973–1980.
- Flinker A, Korzeniewska A, Shestyuk AY, Franaszczuk PJ, Dronkers NF, Knight RT, Crone NE. 2015. Redefining the role of Broca's area in speech. *Proc Natl Acad Sci USA*. 112:2871–2875.
- Frost MA, Goebel R. 2012. Measuring structural–functional correspondence: spatial variability of specialised brain regions after macro-anatomical alignment. *Neuroimage*. 59:1369–1381.
- Gelfand JR, Bookheimer SY. 2003. Dissociating neural mechanisms of temporal sequencing and processing phonemes. *Neuron*. 38:831–842.
- Goldrick MA. 2014. Phonological processing: the retrieval and encoding of word form information in speech production. In: Goldrick MA, Ferreira V, Miozzo M, editors. *The Oxford handbook of language production*. New York: Oxford University Press. p. 228–244.
- Goodale MA. 1988. Hemispheric differences in motor control. *Behav Brain Res*. 30:203–214.
- Grabski K, Lamalle L, Sato M. 2012. Somatosensory-motor adaptation of orofacial actions in posterior parietal and ventral premotor cortices. *PLoS One*. 7:e49117.
- Graff-Radford J, Jones DT, Strand EA, Rabinstein AA, Duffy JR, Josephs KA. 2014. The neuroanatomy of pure apraxia of speech in stroke. *Brain Lang*. 129:43–46.
- Guenther FH. 2006. Cortical interactions underlying the production of speech sounds. *J Commun Disord*. 39:350–365.
- Guenther FH. 2016. *Neural control of speech*. MIT Press, Cambridge, MA.
- Guenther FH, Ghosh SS, Tourville JA. 2006. Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain Lang*. 96:280–301.
- Hesselmann V, Sorger B, Lasek K, Guntinas-Lichius O, Krug B, Sturm V, Goebel R, Lackner K. 2004. Discriminating the cortical representation sites of tongue and lip movement by functional MRI. *Brain Topogr*. 16:159–167.
- Hickok G. 2009. The functional neuroanatomy of language. *Phys Life Rev*. 6:121–143.
- Hickok G. 2014. The architecture of speech production and the role of the phoneme in speech processing. *Lang Cogn Process*. 29:2–20.
- Hickok G, Poeppel D. 2004. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*. 92:67–99.
- Hickok G, Rogalsky C, Chen R, Herskovits EH, Townsley S, Hillis AE. 2014. Partially overlapping sensorimotor networks underlie speech praxis and verbal short-term memory: evidence from apraxia of speech following acute stroke. *Front Hum Neurosci*. 8:649.
- Hillis AE, Work M, Barker PB, Jacobs MA, Breese EL, Maurer K. 2004. Re-examining the brain regions crucial for orchestrating speech articulation. *Brain*. 127:1479–1487.
- Hugdahl K, Raichle ME, Mitra A, Specht K. 2015. On the existence of a generalized non-specific task-dependent network. *Front Hum Neurosci*. 9:430.
- Humbert IA, Robbins J. 2007. Normal swallowing and functional magnetic resonance imaging: a systematic review. *Dysphagia*. 22:266–275.
- Itabashi R, Nishio Y, Kataoka Y, Yazawa Y, Furui E, Matsuda M, Mori E. 2016. Damage to the left precentral gyrus is associated with apraxia of speech in acute stroke. *Stroke*. 47:31–36.
- Jonas S. 1981. The supplementary motor region and speech emission. *J Commun Disord*. 14:349–373.
- Josephs KA, Duffy JR, Strand EA, Machulda MM, Senjem ML, Master AV, Lowe VJ, Jack CR, Whitwell JL. 2012. Characterizing a neurodegenerative syndrome: primary progressive apraxia of speech. *Brain*. 135:1522–1536.
- Josephs KA, Duffy JR, Strand EA, Whitwell JL, Layton KF, Parisi JE, Hauser MF, Witte RJ, Boeve BF, Knopman DS, et al. 2006. Clinicopathological and imaging correlates of progressive aphasia and apraxia of speech. *Brain*. 129:1385–1398.
- Juch H, Zimine I, Seghier ML, Lazeyras F, Fasel JH. 2005. Anatomical variability of the lateral frontal lobe surface: implication for intersubject variability in language neuroimaging. *Neuroimage*. 24:504–514.
- Julian J, Fedorenko E, Webster J, Kanwisher N. 2012. An algorithmic method for functionally defining regions of interest in the ventral visual pathway. *Neuroimage*. 60:2357–2364.
- Kertesz A. 1984. Subcortical lesions and verbal apraxia. In: Rosenbeck JC, McNeil MR, Aronson AE, editors. *Apraxia of speech: physiology, acoustics, linguistics, management*. San Diego, CA: College-Hill Press. p. 73–90.
- Kimura D. 1982. Left-hemisphere control of oral and brachial movements and their relation to communication. *Phil Trans R Soc Lond B Biol Sci*. 298:135–149.
- Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI. 2009. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci*. 12:535–540.
- Kwon M, Lee J-H, Oh JS, Koh JY. 2013. Isolated buccofacial apraxia subsequent to a left ventral premotor cortex infarction. *Neurology*. 80:2166–2167.
- Long MA, Katlowitz KA, Svirsky MA, Clary RC, Byun TM, Majaj N, Oya H, Howard MA, Greenlee JD. 2016. Functional segregation of cortical regions underlying speech timing and articulation. *Neuron*. 89:1187–1193.
- Lotze M, Seggewies G, Erb M, Grodd W, Birbaumer N. 2000. The representation of articulation in the primary sensorimotor cortex. *Neuroreport*. 11:2985–2989.
- Loucks TM, Poletto CJ, Simonyan K, Reynolds CL, Ludlow CL. 2007. Human brain activation during phonation and exhalation: common volitional control for two upper airway functions. *Neuroimage*. 36:131–143.
- Mariën P, van Dun K, Verhoeven J. 2015. Cerebellum and apraxia. *Cerebellum*. 14:39–42.
- Mckay LC, Evans KC, Frackowiak RS, Corfield DR. 2003. Neural correlates of voluntary breathing in humans. *J Appl Physiol*. 95:1170–1178.
- Mesulam M-M. 1998. From sensation to cognition. *Brain*. 121:1013–1052.

- Murphy K, Corfield D, Guz A, Fink G, Wise R, Harrison J, Adams L. 1997. Cerebral areas associated with motor control of speech in humans. *J Appl Physiol*. 83:1438–1447.
- Nieto-Castañón A, Fedorenko E. 2012. Subject-specific functional localizers increase sensitivity and functional resolution of multi-subject analyses. *Neuroimage*. 63:1646–1669.
- Ogar J, Slama H, Dronkers N, Amici S, Gorno-Tempini ML. 2005. Apraxia of speech: an overview. *Neurocase*. 11:427–432.
- Ogar J, Willock S, Baldo J, Wilkins D, Ludy C, Dronkers N. 2006. Clinical and anatomical correlates of apraxia of speech. *Brain Lang*. 97:343–350.
- Pai M-C. 1999. Supplementary motor area aphasia: a case report. *Clin Neurol Neurosurg*. 101:29–32.
- Papoutsis M, de Zwart JA, Jansma JM, Pickering MJ, Bednar JA, Horwitz B. 2009. From phonemes to articulatory codes: an fMRI study of the role of Broca's area in speech production. *Cereb Cortex*. 19:2156–2165.
- Parvizi J, Jacques C, Foster BL, Withoft N, Rangarajan V, Weiner KS, Grill-Spector K. 2012. Electrical stimulation of human fusiform face-selective regions distorts face perception. *J Neurosci*. 32:14915–14920.
- Penfield W, Rasmussen T. 1950. *The cerebral cortex of man: a clinical study of localization of function*. New York: Macmillan.
- Perkell JS. 2012. Movement goals and feedback and feedforward control mechanisms in speech production. *J Neurolinguistics*. 25:382–407.
- Poldrack RA. 2011. Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. *Neuron*. 72:692–697.
- Price CJ. 2012. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage*. 62:816–847.
- Pulvermüller F, Huss M, Kherif F, del Prado Martin FM, Hauk O, Shtyrov Y. 2006. Motor cortex maps articulatory features of speech sounds. *Proc Natl Acad Sci USA*. 103:7865–7870.
- Rauschecker JP, Scott SK. 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci*. 12:718–724.
- Richardson JD, Fillmore P, Rorden C, Lapointe LL, Fridriksson J. 2012. Re-establishing Broca's initial findings. *Brain Lang*. 123:125–130.
- Robin DA, Jacks A, Hageman C, Clark HM, Woodworth G. 2008. Visuomotor tracking abilities of speakers with apraxia of speech or conduction aphasia. *Brain Lang*. 106:98–106.
- Romani C, Galluzzi C. 2005. Effects of syllabic complexity in predicting accuracy of repetition and direction of errors in patients with articulatory and phonological difficulties. *Cogn Neuropsychol*. 22:817–850.
- Romani C, Olson A, Semenza C, Granà A. 2002. Patterns of phonological errors as a function of a phonological versus an articulatory locus of impairment. *Cortex*. 38:541–567.
- Rosenbek J, Kent R, Lapointe L. 1984. Apraxia of speech: an overview and some perspectives. In: Rosenbek JC, McNeil MR, Aronson AE, editors. *Apraxia of speech: physiology, acoustics, linguistics, management*. San Diego, CA: College-Hill Press. p. 1–72.
- Saxe R, Brett M, Kanwisher N. 2006. Divide and conquer: a defense of functional localizers. *Neuroimage*. 30:1088–1096.
- Segawa JA, Tourville JA, Beal DS, Guenther FH. 2015. The neural correlates of speech motor sequence learning. *J Cogn Neurosci*. 27:819–831.
- Shuster LI, Lemieux SK. 2005. An fMRI investigation of covertly and overtly produced mono- and multisyllabic words. *Brain Lang*. 93:20–31.
- Sörös P, Sokoloff LG, Bose A, McIntosh AR, Graham SJ, Stuss DT. 2006. Clustered functional MRI of overt speech production. *Neuroimage*. 32:376–387.
- Staiger A, Ziegler W. 2008. Syllable frequency and syllable structure in the spontaneous speech production of patients with apraxia of speech. *Aphasiology*. 22:1201–1215.
- Strand EA, Duffy JR, Clark HM, Josephs K. 2014. The apraxia of speech rating scale: a tool for diagnosis and description of apraxia of speech. *J Commun Disord*. 51:43–50.
- Takai O, Brown S, Liotti M. 2010. Representation of the speech effectors in the human motor cortex: somatotopy or overlap? *Brain Lang*. 113:39–44.
- Tate MC, Herbet G, Moritz-Gasser S, Tate JE, Duffau H. 2014. Probabilistic map of critical functional regions of the human cerebral cortex: Broca's area revisited. *Brain*. 137:2773–2782.
- Tomaiuolo F, MacDonald J, Caramanos Z, Posner G, Chiavaras M, Evans AC, Petrides M. 1999. Morphology, morphometry and probability mapping of the pars opercularis of the inferior frontal gyrus: an in vivo MRI analysis. *Eur J Neurosci*. 11:3033–3046.
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M. 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*. 15:273–289.
- Uddén J, Bahlmann J. 2012. A rostro-caudal gradient of structured sequence processing in the left inferior frontal gyrus. *Phil Trans R Soc Lond B Biol Sci*. 367:2023–2032.
- Vul E, Hanus D, Kanwisher N. 2009. Attention as inference: selection is probabilistic; responses are all-or-none samples. *J Exp Psychol Gen*. 138:546.
- Whiteside SP, Dyson L, Cowell PE, Varley RA. 2015. The relationship between apraxia of speech and oral apraxia: association or dissociation? *Arch Clin Neuropsychol*. 30:670–682.
- Whitwell JL, Duffy JR, Strand EA, Xia R, Mandrekar J, Machulda MM, Senjem ML, Lowe VJ, Jack CR, Josephs KA. 2013. Distinct regional anatomic and functional correlates of neurodegenerative apraxia of speech and aphasia: an MRI and FDG-PET study. *Brain Lang*. 125:245–252.
- Wildgruber D, Ackermann H, Klose U, Kardatzki B, Grodd W. 1996. Functional lateralization of speech production at primary motor cortex: a fMRI study. *Neuroreport*. 7:2791–2796.
- Wise R, Greene J, Büchel C, Scott SK. 1999. Brain regions involved in articulation. *Lancet*. 353:1057–1061.
- Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. 2011. Large-scale automated synthesis of human functional neuroimaging data. *Nat Methods*. 8:665–670.
- Ziegler W. 2003a. Speech motor control is task-specific: evidence from dysarthria and apraxia of speech. *Aphasiology*. 17:3–36.
- Ziegler W. 2003b. To speak or not to speak: distinctions between speech and nonspeech motor control. *Aphasiology*. 17:99–105.
- Ziegler W. 2006. Distinctions between speech and non-speech motor control: a neurophonetic view. In: Tabain M, Harrington J, editors. *Speech production: models, phonetic processes, and techniques*. New York: Psychology Press. p. 41–54.
- Ziegler W. 2016. The phonetic cerebellum: cerebellar involvement in speech sound production. In: Marien P, Manto M, editors. *The linguistic cerebellum*. Boston: Elsevier. p. 1–32.
- Ziegler W, Ackermann H. 2014. Neural bases of phonological and articulatory processing. In: Goldrick MA, Ferreira V, Miozzo M, editors. *The Oxford handbook of language production*. New York: Oxford University Press. p. 275–291.