

## Research Article

# The Neural Circuitry Underlying the “Rhythm Effect” in Stuttering

Saul A. Frankford,<sup>a</sup> Elizabeth S. Heller Murray,<sup>a</sup> Matthew Masapollo,<sup>a</sup> Shanjing Cai,<sup>a</sup> Jason A. Tourville,<sup>a</sup> Alfonso Nieto-Castañón,<sup>a</sup> and Frank H. Guenther<sup>a,b,c,d</sup>

**Purpose:** Stuttering is characterized by intermittent speech disfluencies, which are dramatically reduced when speakers synchronize their speech with a steady beat. The goal of this study was to characterize the neural underpinnings of this phenomenon using functional magnetic resonance imaging.

**Method:** Data were collected from 16 adults who stutter and 17 adults who do not stutter while they read sentences aloud either in a normal, self-paced fashion or paced by the beat of a series of isochronous tones (“rhythmic”). Task activation and task-based functional connectivity analyses were carried out to compare neural responses between speaking conditions and groups after controlling for speaking rate.

**Results:** Adults who stutter produced fewer disfluent trials in the rhythmic condition than in the normal condition. Adults who stutter did not have any significant changes in activation between the rhythmic condition and the normal condition, but when groups were collapsed, participants had greater

activation in the rhythmic condition in regions associated with speech sequencing, sensory feedback control, and timing perception. Adults who stutter also demonstrated increased functional connectivity among cerebellar regions during rhythmic speech as compared to normal speech and decreased connectivity between the left inferior cerebellum and the left prefrontal cortex.

**Conclusions:** Modulation of connectivity in the cerebellum and prefrontal cortex during rhythmic speech suggests that this fluency-inducing technique activates a compensatory timing system in the cerebellum and potentially modulates top-down motor control and attentional systems. These findings corroborate previous work associating the cerebellum with fluency in adults who stutter and indicate that the cerebellum may be targeted to enhance future therapeutic interventions.

**Supplemental Material:** <https://doi.org/10.23641/asha.14417681>

Stuttering is a speech disorder that impacts the production of smooth and timely articulations of planned utterances. Stuttering typically emerges early in childhood and persists over the life span for 1% of the population (Craig et al., 2009; Yairi & Ambrose, 1999). Speech of people who stutter (PWS) is characterized by perceptually salient repetitions and prolongations of individual phonemes, as well as abnormal silent pauses at the onset of syllables

and words accompanied by tension in the articulatory musculature (Max, 2004). These disfluencies are often accompanied by other secondary behaviors, such as eye blinking and facial grimacing (Guitar, 2014). Along with these more overt characteristics, stuttering also has a severe impact on those who experience it, including increased social anxiety and decreased self-confidence, emotional functioning, and overall mental health (Craig et al., 2009; Craig & Tran, 2006, 2014). Gaining a better understanding of how and why stuttering occurs will help to lead to more targeted therapies and improve quality of life for PWS.

Considerable effort has been made to identify the core pathology underlying stuttering (for reviews, see Max, 2004; Max et al., 2004). More recently, diverse brain imaging modalities have been used to examine how the brains of PWS differ from those who do not and how these measures change in different speaking scenarios or following therapy (see Etchell et al., 2018, for a complete literature review). Studies have consistently found that PWS show structural and functional differences in the brain network pertaining to speech initiation and timing (cortico-thalamo-basal ganglia

<sup>a</sup>Department of Speech, Language & Hearing Sciences, Boston University, MA

<sup>b</sup>Department of Biomedical Engineering, Boston University, MA

<sup>c</sup>Department of Radiology, Massachusetts General Hospital, Boston

<sup>d</sup>The Picower Institute for Learning and Memory, Massachusetts Institute of Technology, Cambridge

Correspondence to Saul A. Frankford: [sfrankford@meei.harvard.edu](mailto:sfrankford@meei.harvard.edu)

Editor-in-Chief: Cara E. Stepp

Editor: Bridget Walsh

Received June 8, 2020

Revision received December 23, 2020

Accepted January 12, 2021

[https://doi.org/10.1044/2021\\_JSLHR-20-00328](https://doi.org/10.1044/2021_JSLHR-20-00328)

**Publisher Note:** This article is part of the Special Issue: Selected Papers From the 2020 Conference on Motor Speech—Basic Science and Clinical Innovation.

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

motor loop; Chang & Zhu, 2013; Giraud, 2008; Lu, Peng, et al., 2010) and reduced structural integrity in speech planning areas (left ventral premotor cortex [vPMC] and inferior frontal gyrus [IFG]; Beal et al., 2013, 2015; Chang et al., 2008, 2011; Garnett et al., 2018; Kell et al., 2009; Lu et al., 2012). Functionally, previous work has indicated that, during speech, adults who stutter (AWS) have reduced activation in left hemisphere auditory areas (Belyk et al., 2015; Braun et al., 1997; Chang et al., 2009; De Nil et al., 2008, 2000; Fox et al., 1996; Van Borsel et al., 2003) and overactivation in right hemisphere structures (Braun et al., 1997; De Nil et al., 2000; Fox et al., 1996, 2000; Ingham et al., 2000; Van Borsel et al., 2003), which are typically nondominant for language processing. These studies strongly suggest that stuttering occurs as the result of impaired speech timing, planning, and auditory processing and that brain structures not normally involved in speech production are potentially recruited to compensate.

In addition to these task activation analyses, previous studies have examined task-based functional connectivity (i.e., activation coupling between multiple brain areas during a speaking task) differences between AWS and adults who do not stutter (ANS). Some studies show reduced connectivity between the left IFG and the left precentral gyrus in AWS (Chang et al., 2011; Lu et al., 2009), which suggests an impairment in translating speech plans for motor execution (Guenther, 2016). Other studies show group differences in connectivity between auditory, motor, premotor, and subcortical areas (Chang et al., 2011; Kell et al., 2018; Lu, Chen, et al., 2010; Lu et al., 2009; Lu, Peng, et al., 2010). Results of these task-based connectivity studies, as well as resting-state and structural connectivity studies (e.g., Chang & Zhu, 2013; Sitek et al., 2016), have made it apparent that stuttering behavior is not merely the result of disruptions to one or more separate brain regions, but also differences in the ability for brain regions to communicate with one another during speech.

Beyond examining neural activation in AWS during typical speech, imaging studies have also looked at activation during conditions where AWS speak more fluently. One such condition that has been widely examined behaviorally is the *rhythm effect* in which stuttering disfluencies are dramatically reduced when speakers synchronize their speech movements with isochronous pacing stimuli (Azrin et al., 1968; Barber, 1940; Hutchinson & Norris, 1977; Stager et al., 1997; Toyomura et al., 2011). These fluency-enhancing effects are robust; they occur regardless of whether the pacing stimulus is presented in the acoustic or visual modalities (Barber, 1940), can be induced even by an imagined rhythm (Barber, 1940; Stager et al., 2003), and occur independently of speaking rate (Davidow, 2014; Hanna & Morris, 1977). Previous studies investigating changes in brain activation during the rhythm effect (Braun et al., 1997; Stager et al., 2003; Toyomura et al., 2011, 2015) have found that, during isochronous speech, both AWS and ANS had increased activation in speech-related auditory and motor regions of cortex as well as parts of the basal ganglia. These activation increases were especially pronounced for AWS as compared to ANS. Toyomura et al. (2011) also demonstrated

that these activation increases occurred in regions displaying underactivation during the unpaced speaking condition. This suggests that pacing speech, along with a metronome, improves fluency by “normalizing” underactivation in speech production regions. In light of the functional connectivity studies mentioned previously, characterizing changes in brain connectivity between typical and isochronously paced speech could illuminate how external pacing leads to normalized activation in the speech network and, ultimately, fluency.

In this study, we employed functional magnetic resonance imaging (fMRI) during an overt isochronously paced sentence-reading task in AWS and ANS to characterize modulation of brain activation and functional connectivity related to the rhythm effect in stuttering. In addition, this study sought to address an important issue not previously accounted for in neuroimaging studies of the rhythm effect: a reduced speaking rate in the paced compared to the unpaced condition. Reduced speaking rate and paced speech can both induce fluency in AWS (Andrews et al., 1982), but the effects are dissociable—the rhythm effect increases fluency even when speaking rates are matched between speaking conditions (Davidow, 2014). Since brain activation is also modulated by speaking rate (Fox et al., 2000; Riecker et al., 2006), activation changes between paced and unpaced conditions may reflect either the planning/production features or the fluency-inducing effect of both, unless rate is accounted for. Two prior studies (Braun et al., 1997; Stager et al., 2003) examined general differences between “fluent” and “disfluent” speaking conditions, aiming to characterize the neural underpinnings of fluency without controlling for features that contributed (e.g., rate, speaking style, percent voicing). Toyomura et al. (2011) attempted to control for rate differences between the conditions by instructing participants to speak at similar rates during both conditions. However, they still found a significantly reduced speaking rate in the metronome-paced condition that was not accounted for in their analyses. Separating out the effects of rate would help elucidate the neural underpinnings of the rhythm effect itself. In this study, a combination of training and analysis procedures was used to accomplish this.

## Method

The current study complied with the principles of research involving human subjects as stipulated by the Boston University (BU) Institutional Review Board (Protocol 2421E) and the Massachusetts General Hospital (MGH) Human Research Committee, and participants gave informed consent before taking part. The entire experimental procedure took approximately 2 hr, and subjects received monetary compensation.

## Subjects

Sixteen AWS (11 men and five women, aged 18–58 years,  $M_{\text{age}} = 29.9$  years,  $SD = 12.9$ ) and 17 ANS (11 men and six women, aged 18–49 years,  $M_{\text{age}} = 28.7$  years,  $SD = 8.1$ ) from the greater Boston area were included in the final

analyses. Age was not significantly different between groups (two-sample *t* test,  $t = 0.31$ ,  $p = .756$ ). Subjects were native speakers of American English who reported normal (or corrected-to-normal) vision and no history of hearing, speech, language, or neurological disorders (aside from persistent developmental stuttering for the AWS). Handedness was measured with the Edinburgh Handedness Inventory (Oldfield, 1971). Using this metric, all AWS were found to be right-handed (scoring greater than 40), but there was more variability among ANS (13 right-handed, one left-handed, and three ambidextrous). There was a significant difference in handedness score between groups (Wilcoxon rank-sum test,  $z = 2.29$ ,  $p = .022$ ); therefore, handedness score was included as a covariate in all group imaging comparisons. For each stuttering participant, stuttering severity was determined using the Stuttering Severity Instrument–Fourth Edition (SSI-4; Riley, 2008; mean score = 23.1, range: 9–42; see Table 1 for individual participants). Four additional subjects (three AWS and one ANS) were also tested, but they were excluded during data inspection (described below in the Behavioral Analysis and Task Activation fMRI Analysis sections).

### fMRI Paradigm

Sixteen 8-syllable sentences were selected from the Revised List of Phonetically Balanced Sentences (Harvard Sentences; Institute of Electrical and Electronics Engineers, 1969; see Appendix A). These sentences, composed of one- and two-syllable words, contain a broad distribution of English speech sounds (e.g., “The juice of lemons makes fine punch”). During a functional brain-imaging session, subjects read aloud the stimulus sentences under two different speaking conditions, one in which individual syllables were paced by isochronous auditory beats (i.e., the *rhythm* condition) and one in which syllables were not paced (i.e., the *normal* condition). For each

trial, subjects were presented with eight isochronous tones (1000 Hz, 25-ms duration), with a 270ms interstimulus interval. This resulting rate of approximately 222 beats per minute was chosen so that participants’ speech would approximate the rate of the normal condition (based on previous estimates of mean speaking rate in English; Davidow, 2014; Pellegrino et al., 2004). Participants were instructed to refrain from using any part of their body (e.g., finger or foot) to tap to the rhythm.

To avoid confounding the auditory region blood oxygen level–dependent (BOLD) response to the pace tone and speech auditory feedback, the pacing tones were terminated prior to the presentation of the orthographic stimulus. During a rhythm or normal trial, the orthography of a given sentence was presented with the corresponding trial identifier (i.e., “rhythm” or “normal”) presented above the sentence. From this identifier, subjects were instructed to either read the sentence “in a rhythmic way” by aligning each syllable to a beat or in a natural way. Thus, on rhythm trials, subjects used the tones to pace their forthcoming speech, while on normal trials, they read the stimuli at a normal speaking rate, rhythm, and intonation (see Appendix B for detailed instructions). The font color was either blue for rhythm or green for normal or vice versa, and colors were counter-balanced across subjects. Subjects were instructed to begin reading aloud immediately after the sentence appeared on the screen. In the event that they made a mistake, they were asked to refrain from producing any corrections and remain silent until the next trial. Silent baseline trials were also included wherein subjects heard the tones and saw a random series of typographical symbols (e.g., “+ \ ^ & \$ / | \ \$ = [ ] \* % / - @ \ | - % - /”) clustered into wordlike groupings (matched to stimulus sentences); subjects refrained from speaking during these trials.

Subjects participated in a behavioral experiment (not reported here) prior to the imaging experiment that gave

**Table 1.** Demographic and stuttering severity data from adults who stutter (AWS).

Subject ID	Age	Gender	SSI-4 composite	SSI-Mod	Disfluency rate
AWS01	19	F	28	19	0%
AWS02	22	F	31	26	3.03%
AWS03	31	F	30	22	3.03%
AWS04	21	M	9	7	1.92%
AWS05	58	M	14	11	0%
AWS06	23	M	42	29	0%
AWS07	53	M	27	22	0%
AWS08	44	M	20	16	0%
AWS09	20	M	18	15	1.52%
AWS10	22	M	27	18	3.02%
AWS11	21	M	19	16	6.06%
AWS12	20	M	24	14	1.52%
AWS13	18	F	14	11	0%
AWS14	35	M	30	19	0%
AWS15	42	M	22	17	1.52%
AWS16	29	M	14	12	0%

*Note.* SSI-4 = Stuttering Severity Index–Fourth Edition; SSI-Mod = a modified version of the SSI-4 that does not include a subscore related to concomitant movements; Disfluency rate = the percentage of trials containing disfluencies during the normal speech condition; F = female; M = male.

them experience with the speech stimuli and the task. The time between this prior exposure and the present experiment ranged from 0 to 424 days. Immediately prior to the imaging session, subjects practiced each sentence under both conditions until they demonstrated competence with the task and sentence production. Subjects also completed a set of six practice trials in the scanner prior to fMRI data collection. To control basic speech parameters across conditions and groups, subjects were provided with performance feedback on their overall speech rate and loudness during practice only. Following this practice set, subjects completed between two and four experimental runs of test trials, depending on time constraints (14 ANS and 14 AWS completed four runs, three ANS and one AWS completed three runs, and one AWS completed two runs). During the experimental session, verbal feedback was provided between runs if subjects consistently performed outside the specified speech rate (mean syllable duration = 220–320 ms). Each run consisted of 16 rhythm trials, 16 normal trials, and 16 baseline trials, pseudorandomly interleaved within each run for each subject. All trials were audio-recorded for later processing.

### **Data Acquisition**

MRI data for this study were collected at two locations: the Athinoula A. Martinos Center for Biomedical Imaging at the MGH, Charlestown Campus (nine AWS, nine ANS) and the Cognitive Neuroimaging Center at BU (eight AWS, eight ANS). At MGH, images were acquired with a 3T Siemens Skyra scanner and a 32-channel head coil, while a 3T Siemens Prisma scanner with a 64-channel head coil was used at BU. At each location, subjects lay supine in the scanner, and functional volumes were collected using a gradient echo, echo-planar imaging BOLD sequence (repetition time [TR] = 11.5 s, acquisition time = 2.47 s, echo time = 30 ms, flip angle = 90°). Each functional volume covered the entire brain and was composed of 46 axial slices (64 × 64 matrix) acquired in interleaved order and accelerated using a simultaneous multislice factor of 3, with a 192-mm field of view. The in-plane resolution was 3.0 × 3.0 mm<sup>2</sup>, and slice thickness was 3.0 mm with no gap. Additionally, a high-resolution T1-weighted whole-brain structural image was collected from each participant to anatomically localize the functional data (MPRAGE [magnetization-prepared rapid gradient-echo] sequence, 256 × 256 × 176 mm<sup>3</sup> volume, with a 1-mm isotropic resolution, TR = 2.53 s, inversion time = 1,100 ms, echo time = 1.69 ms, flip angle = 7°).

Functional data were acquired using a sparse image acquisition paradigm (Eden et al., 1999; Hall et al., 1999) that allowed participants to produce the target sentences during silent intervals between volume acquisitions. Volumes were acquired 5.7–8.17 s after visual stimulus presentation to ensure a 4- to 6-s delay between the middle of sentence production (approximately 2.3 s post-sentence presentation) and the middle of the acquisition (approximately 6.9 s post-sentence presentation), aligning the acquisition to the peak of the canonical task-related BOLD response to the subject's production (Poldrack et al., 2011). Prior work has shown there is variation

in the timing of this hemodynamic response across tasks, brain regions, and participants (Handwerker et al., 2004; Janssen & Mendieta, 2020). However, since the functional volumes are acquired over 2.47 s, sentences are produced over the course of about 2 s, and there is a random amount of jitter between the start of the sentence production and the start of the acquisition at each trial; the single acquisition provides a broad sampling of the hemodynamic response across a range of different delay times. Furthermore, by scanning after speech production has ended, this paradigm reduces head motion-induced scan artifacts, eliminates the influence of scanner noise on speaker performance, and allows subjects to perceive their own self-generated auditory feedback in the absence of scanner noise (e.g., Gracco et al., 2005). A schematic representation of the trial structure and timeline is shown in Figure 1.

Visual stimuli were projected onto a screen viewed from within the scanner via a mirror attached to the head coil. Auditory stimuli were delivered to both ears through Sennheiser Model S-14 MRI-compatible earphones using MATLAB (The MathWorks). Subjects' utterances were transduced with a Fibersound Model FOM1-MR-30m fiberoptic microphone, sent to a laptop (Lenovo ThinkPad W540), and recorded using MATLAB. Subjects took a short break after completing each run.

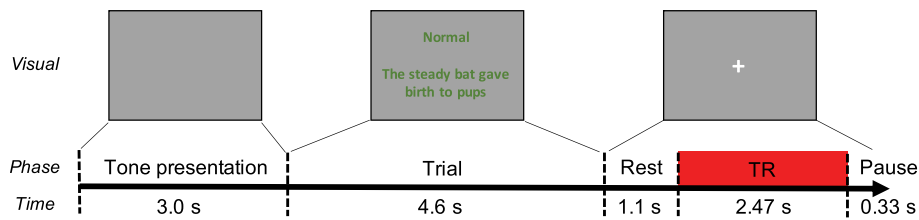
### **Behavioral Analysis**

The open-source large-vocabulary continuous speech recognition engine Julius (Lee & Kawahara, 2009) was used in conjunction with the free VoxForge American English acoustic models (voxforge.org) to perform phoneme-level alignment on the sentence recordings. This resulted in phoneme boundary timing information for every trial. A researcher manually inspected each trial to ensure correct automatic detection of phoneme boundaries. Any trials in which the subject made a reading error, a condition error (i.e., spoke at an isochronous pace when they were cued to speak normally or vice versa), or a disfluency categorized as a stutter by a licensed speech-language pathologist were eliminated from further behavioral analysis. One ANS who made consistent condition errors was eliminated from further analysis. One AWS was eliminated from further analysis due to an insufficient number of fluent trials during the normal speech condition (six of 64 attempted). Neither were included in the total participant count in the Subjects section.

To evaluate whether there was a fluency-enhancing effect of isochronous pacing, the percentage of trials eliminated due to stuttering in the AWS group was compared between the two speaking conditions using a nonparametric Wilcoxon signed-ranks test. Measures of the total sentence duration and intervocalic timing from each trial were also extracted to determine the rate and isochronicity of each production. Within a sentence, the average time between the centers of the eight successive vowels was calculated to determine the intervocalic interval (IVI). The reciprocal (1/IVI) was then calculated, resulting in a measure of speaking rate in units of IVIs per second. The coefficient of variation for IVIs (CV-IVIs) was also calculated by dividing the



**Figure 1.** Schematic diagram illustrating the temporal structure of stimulus presentation during functional data acquisition. At the start of each trial, isochronous tone sequences were presented for 3.0 s. The visual stimulus then appeared and remained on screen for 4.6 s. At 1.1 s after stimulus offset, a whole-brain volume was acquired. The next trial started 0.33 s after data acquisition was complete. TR = repetition time.



standard deviation of IVIs by the mean IVI. A higher CV-IVI indicates higher variability of IVI, while a CV-IVI of 0 reflects perfect isochronicity. Rate and CV-IVI were compared between groups and conditions using a mixed-design analysis of variance. A Bonferroni correction was applied across these two analyses to account for multiple testing.

### Task Activation fMRI Analysis

#### Preprocessing

Following data collection, all images were processed through two preprocessing pipelines: a surface-based pipeline for cortical activation analyses and a volume-based pipeline for subcortical and cerebellar analyses. For both the surface- and volume-based pipelines, functional images from each subject were simultaneously realigned to the mean subject image and unwrapped (motion-by-inhomogeneity interactions) using SPM12's realign and unwarp procedure (Andersson et al., 2001). Outlier scans were detected with artifact detection tools ([https://www.nitrc.org/projects/artifact\\_detect/](https://www.nitrc.org/projects/artifact_detect/)) based on motion displacement (scan-to-scan motion threshold of 0.9 mm) and mean signal change (scan-to-scan signal change threshold of 5 SDs above the mean). For the surface-based pipeline, functional images from each subject were then coregistered with their high-resolution T1 structural images and resliced using SPM12's intermodal registration procedure with a normalized mutual information objective function. The structural images were segmented into white matter, gray matter, and cerebrospinal fluid, and cortical surfaces were reconstructed using the FreeSurfer image analysis suite (freesurfer.net; Fischl et al., 1999). Functional data were then resampled at the location of the FreeSurfer *fsaverage* tessellation of each subject-specific cortical surface. For the vertex-level analyses (see Second-Level Group Analyses section), surfaces were additionally smoothed using iterative diffusion smoothing with 40 diffusion steps (equivalent to an 8-mm full-width half-maximum smoothing kernel; Hagler et al., 2006).

For the volume-based pipeline, after the outlier detection step, functional volumes were then simultaneously segmented and normalized directly to Montréal Neurological Institute (MNI) space using SPM12's combined normalization and segmentation procedure (Ashburner & Friston, 2005). For the voxel-level analyses (see Second-Level Group Analyses section), volumes were also smoothed using an 8-mm full-width half-maximum smoothing kernel. A mask was then

applied, such that only voxels within the subcortical structures were submitted to subsequent analyses. The original T1 structural image from each subject was also centered, segmented, and normalized using SPM12. Following preprocessing, two AWS (not included in the 16 described in the Subjects section) were eliminated from subsequent analyses: one due to excessive head motion in the scanner (> 1.5 mm average scan-to-scan motion) and one due to structural brain abnormalities.

#### First-Level Analysis

After preprocessing, BOLD responses were estimated for each subject using a general linear model (GLM) in SPM12. Because images were collected in a sparse sequence with a relatively long TR, the BOLD response for each trial (event) was modeled as an individual epoch. The model included regressors for each of the conditions of interest: normal, rhythm, and baseline. Trials that contained reading errors, condition errors, or disfluencies were modeled as a single separate condition of noninterest. To control for differences in rate between the two conditions (see the Results section), trial-by-trial mean IVIs were centered and added as a covariate of noninterest. These regressors were collapsed across runs to maximize power while controlling for potential differences in the number of trials produced without errors or disfluencies. For each run, regressors were added to remove linear effects of time (e.g., signal drift, adaptation) in addition to six motion covariates (taken from the realignment step) and a constant term, as well as outlier regressors (one regressor per identified outlier) to remove the effects of acquisitions with excessive scan-to-scan motion or global signal change (estimated from the artifact detection step, described above). The first-level GLM regressor coefficients for the three conditions of interest were estimated at each surface vertex and subcortical voxel. The mean normal speech and rhythm speech coefficients were then contrasted with the baseline condition to yield contrast effect size values for the two contrasts of interest (normal–baseline and rhythm–baseline).

#### Region of Interest Definition

Cortical regions of interest (ROIs) were labeled according to a modified version of the SpeechLabel atlas previously described in Cai et al. (2014); the atlas divides the cortex into macro-anatomically defined ROIs specifically tailored for studies of speech. Labels are applied by mapping the

atlas from the FreeSurfer *fsaverage* cortical surface template to each individual surface reconstruction.

Subcortical and cerebellar ROIs were extracted from multiple atlases. Thalamic ROIs were extracted from the mean atlas of thalamic nuclei described by Krauth et al. (2010). Basal ganglia ROIs were derived from the nonlinear normalized probabilistic atlas of basal ganglia described by Keuken et al. (2014). Each ROI was thresholded at a minimum probability threshold of 33% and combined in a single labeled volume in the atlas's native space (the MNI104 template). Cerebellar ROIs were derived from the SUIT (spatially unbiased infra-tentorial template) 25% maximum probability atlas of cerebellar regions (Diedrichsen, 2006; Diedrichsen et al., 2009, 2011). Each atlas was nonlinearly registered to the SPM12 MNI152 template and then combined into a single labeled volume.

### Second-Level Group Analyses

Group activation differences were examined in the two speech conditions compared to baseline (normal–baseline, rhythm–baseline) as well as the Group  $\times$  Condition interaction. Additionally, differences between the two speech conditions (rhythm–normal) were examined in each group separately. All group-level analyses were performed using a GLM with random effects across subjects. Group comparisons included the following four control covariates: (a) subject motion (average framewise displacement score for each subject), (b) acquisition site (MGH vs. BU), (c) handedness (due to significant difference in handedness between the two groups; see the Subjects section), and (d) stuttering severity within the AWS group only. This severity covariate was a modification of the SSI-4 score, heretofore termed “SSI-Mod.” SSI-Mod removes the secondary concomitants subscore from each subject's SSI-4 score, thus focusing the measure on speech-related function. The SSI-Mod and SSI-4 composite scores for each subject are included in Table 1. With 16 AWS and 17 ANS and four control covariates, power is sufficient (greater than 80%) to detect at a  $p < .05$  false positive control level large between-group differences (Cohen's  $d > 0.87$ ). It is not uncommon to find or expect such large effects in the context of voxel- or surface-level analyses, and these sample sizes are comparable to or larger than those of similar studies (Stager et al., 2003; Toyomura et al., 2011, 2015). Additional regression analyses were carried out to determine whether stuttering severity, measured by the SSI-Mod, or disfluencies occurring during the experiment were correlated with task activation. Because very few disfluencies occurred during the rhythm condition, we were only able to calculate the correlation between the percentage of disfluencies occurring during normal trials (“disfluency rate”) and the normal–baseline activation. Note that because trials containing disfluencies were regressed out of the first-level effects, correlations with disfluency rate are capturing activation related to the *propensity* to stutter and not disfluent speech itself.

Two sets of group-level analyses were carried out to detect activation differences across groups and conditions: analyses at the level of the vertex (cortical) or voxel (subcortical) and exploratory ROI analyses. For the vertex/voxel

analyses, the GLM was carried out on the smoothed data at each unit. Unit-wise statistics were first thresholded at a height threshold of  $p < .01$ , uncorrected. Cluster-level statistics were then estimated using a permutation/randomization analysis with 1,000 simulations (Bullmore et al., 1999), and only clusters below  $p_{\text{FDR}} < .05$  threshold are reported (topological false discovery rate [FDR]; Chumbley et al., 2010). Additional ROI analyses were performed to determine if activation from other brain regions was also modulated by group or condition at a less strict threshold. First-level contrast effects calculated from nonsmoothed data were averaged within each ROI. For each exploratory analysis, ROIs below a  $p < .05$  uncorrected threshold are reported.

### Functional Connectivity Analysis

#### Preprocessing and Analysis

Seed-based functional connectivity analyses were carried out using the CONN toolbox (Whitfield-Gabrieli & Nieto-Castanon, 2012). The same preprocessed data used for the task activation analysis were used for the functional connectivity analysis. The seeds for this analysis comprised a subset of the ROIs used in the exploratory task activation analysis, defined in either *fsaverage* surface (cortical) or MNI volume (subcortical) space. These included regions with significant positive activation (thresholded at one-sided  $p < .05$  and corrected for multiple comparisons using a false discovery rate correction within each contrast; Benjamini & Hochberg, 1995) in the normal–baseline or rhythm–baseline contrasts, or significant rhythm–normal activation in either direction (thresholded at two-sided  $p < .05$ , uncorrected) across all subjects. In addition, prior work has found that connectivity between left orbitofrontal regions and the cerebellum is both increased in adults who have spontaneously recovered from stuttering (Kell et al., 2018) and negatively associated with severity (Sitek et al., 2016), indicating a potential common substrate of fluency in AWS. To determine whether connectivity between these regions is also found in rhythm-induced fluency, three left orbitofrontal regions were added as seeds (see Supplemental Figures S10 and S11 for a complete list).

The BOLD time series was first averaged within seed ROIs. To include connections between the speech production network and other regions that potentially have a moderating effect on this network, the target area in this analysis was extended to the whole brain. The target functional volume data were smoothed using an 8-mm full-width half-maximum Gaussian smoothing kernel. Following preprocessing, an aCompCor (anatomical component-based noise correction; Behzadi et al., 2007) denoising procedure was used to eliminate extraneous motion, physiological, and artifactual effects from the BOLD signal in each subject. In each seed ROI and every voxel in the smoothed brain volume, denoising was carried out using a linear regression model (Nieto-Castañón, 2020) that included five white matter regressors; five cerebrospinal fluid regressors; six subject motion parameters plus their first-order temporal derivatives; scrubbing regressors to remove the effects of outlier scans (from artifact detection, described

above); and separate regressors for each run/session (constant effects and first-order linear trends), task condition (main and first-order derivative terms), and error trial. No bandpass filter was applied in order to preserve high-frequency fluctuations in the residual data.

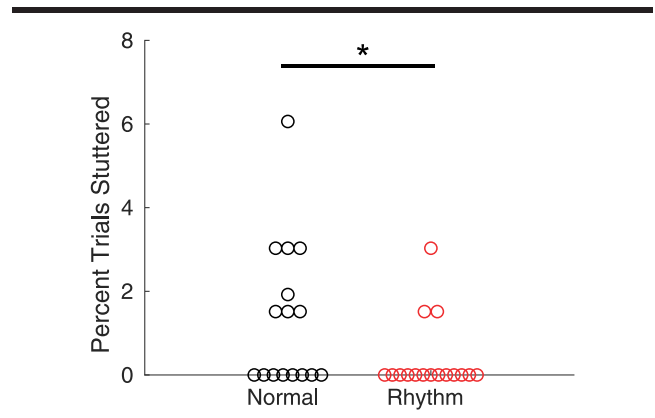
For each participant, a generalized psychophysiological interaction (PPI; McLaren et al., 2012) analysis was implemented using a multiple regression model, predicting the signal in each target voxel with three sets of regressors: (a) the BOLD time series in a seed ROI, characterizing baseline connectivity between a seed ROI and each target voxel; (b) the main effects of each of the task conditions (normal, rhythm, and baseline), characterizing direct functional responses to each task in the target voxel; and (c) their seed-time-series-by-task interactions (PPI terms), characterizing the relative changes in functional connectivity strength associated with each task. The implementation of PPI in CONN used in this article (Nieto-Castañón, 2020) is based on the original Friston et al. (1997) formulation, where the interaction is modeled and estimated at the level of the BOLD signal directly. Among other potential benefits, this allows the direct application of PPI and generalized PPI to the analysis of sparse acquisition data sets. Second-level random effects analyses were then used to compare these interaction terms within and between groups and conditions, specifically the rhythm–normal contrast in AWS and ANS and the Group  $\times$  Condition interaction. Additional analyses examining the correlation between normal–baseline and SSI\_Mod, rhythm–baseline and SSI\_Mod, and normal–baseline and disfluency rate in the normal condition were also carried out. All group-level analyses included the same four control covariates used in the task activation analyses. For each comparison, separate analyses were run from the 116 seed ROIs to the whole brain. Within each analysis, a two-step thresholding procedure was used; voxels were thresholded at a  $p < .001$  height threshold, followed by a cluster size threshold of  $p_{FDR} < .05$  estimated using random Gaussian field theory (Worsley et al., 1996). To control for familywise error (FWE) across the 116 separate seed-to-voxel analyses, a within-comparison Bonferroni correction was applied so that only significant clusters with  $p_{FDR} < .00043$  ( $0.05/116$ ) were reported.

## Results

### Behavioral Analysis

Stuttering occurred infrequently over the course of the experiment, with seven out of 16 AWS producing no disfluencies. There was, however, a significantly lower percentage of disfluent trials in the rhythm condition (0.38%) compared to the normal condition (1.35%;  $W = 42$ ,  $p = .023$ ; see Figure 2). There was no Group  $\times$  Condition interaction or group main effect on speaking rate, but there was a significant main effect of condition with normal trials (3.773 IVI/s) produced at a faster rate than rhythm trials (3.463 IVI/s),  $F(1, 31) = 54.7$ ,  $p_{FWE} < .001$ . To examine

**Figure 2.** Comparison of disfluencies between the normal and rhythm conditions for adults who stutter. Circles represent individual participants. \* $p < .05$ .



whether this reduction in rate led to increased fluency rather than the isochronous pacing, we tested for a correlation between the change in speech rate and the reduction in disfluencies. These two measures were not significantly correlated ( $r = -.07$ ,  $p = .80$ ). For isochronicity, there was no main effect of group or Group  $\times$  Condition interaction. There was a significant main effect of condition, where subjects had a lower CV-IVI (greater isochronicity) in the rhythm condition (0.13) than the normal condition (0.25),  $F(1, 31) = 492.0$ ,  $p_{FWE} < .001$ . For complete results regarding speaking rate and CV-IVI, see Table 2.

### Task Activation fMRI Analysis

For the vertex/voxel-wise analysis, no significant differences were found between groups for either normal–baseline or rhythm–baseline (vertex/voxel-level  $p < .01$ , cluster-level  $p_{FDR} < .05$ ). Similarly, no clusters showed a significant interaction between groups and conditions. Within the AWS group, there were no significant differences between the two conditions. Because there were no significant group differences in either condition and no significant Group  $\times$  Condition interactions, the rhythm–normal analysis was collapsed across groups to improve power. Clusters that had greater activation during the rhythm condition than the normal condition (vertex/voxel-level  $p < .01$ , cluster-level  $p_{FDR} < .05$ ) are shown in Table 3 and Figure 3. These six clusters include the left hemisphere cortex spanning posterior Sylvian fissure (planum temporale [PT] and parietal operculum, supramarginal gyrus [SMg], and intraparietal sulcus), left posterior superior parietal lobule (SPL), left supplementary motor area (SMA), right SPL, right SMg, and right dorsal premotor cortex (dPMC). No regions in the cerebral cortex or subcortical structures were found to be more active during the normal condition than the rhythm condition.

In the exploratory ROI analysis, AWS had increased activation in the left middle temporo-occipital cortex ( $p = .004$ ), left posterior middle temporal gyrus ( $p = .010$ ),

**Table 2.** Descriptive and inferential statistics for speaking rate and coefficient of variation for intervocalic intervals (CV-IVI).

Measure	ANS		AWS		Main effect of group	Main effect of condition	Interaction
	Normal	Rhythm	Normal	Rhythm			
Speaking rate (IVI/s)	3.797 ± 0.086	3.456 ± 0.080	3.748 ± 0.164	3.470 ± 0.173	$F(1, 31) = 0.1,$ $p_{FWE} = 1$	<b><math>F(1, 31) = 54.7,</math></b> <b><math>p_{FWE} &lt; .001</math></b>	$F(1, 31) = 0.6,$ $p_{FWE} = .92$
CV-IVI	0.259 ± 0.013	0.127 ± 0.006	0.251 ± 0.019	0.132 ± 0.007	$F(1, 31) = 0.1,$ $p_{FWE} = 1$	<b><math>F(1, 31) = 492.0,</math></b> <b><math>p_{FWE} &lt; .001</math></b>	$F(1, 31) = 1.4,$ $p_{FWE} = .48$

Note. Error estimates indicate 95% confidence intervals. Significant effects are highlighted in bold. ANS = adults who do not stutter; AWS = adults who stutter; FWE = familywise error rate.

and left anterior ventral superior temporal sulcus ( $p = .042$ ) for the normal–baseline contrast compared to ANS and decreased activation in cerebellar vermis X ( $p = .049$ ; see Supplemental Table S1). In the rhythm–baseline contrast, AWS had reduced activation in the left anterior frontal operculum ( $p < .009$ ), midline cerebellar vermis VIIIb ( $p < .008$ ) and cerebellar vermis VIIIa ( $p < .042$ ), and right anterior middle temporal gyrus ( $p < .040$ ) and cerebellar lobule X ( $p < .046$ ) compared to ANS (see Supplemental Table S1). Also, in this exploratory analysis, interactions were found in a number of cortical and subcortical ROIs, including bilateral auditory regions and left inferior cerebellum (see Supplemental Table S2 and Supplemental Figure S2 for complete results). In all cases, ANS had increased activation in the rhythm condition compared to normal, while AWS showed no change or a decrease. For complete exploratory ROI results for the rhythm–normal analysis in each group separately and combined, see Supplemental Table S3 and Supplemental Figures S3–S5.

### Brain–Behavior Correlation Analyses

In our vertex/voxel-wise analysis, no significant clusters were found showing a correlation between SSI-Mod and normal–baseline or rhythm–baseline, or between disfluency rate and normal–baseline. Exploratory results can be found in Supplemental Table S4 and Supplemental Figures S6–S9.

Of note, positive correlations were found between SSI-Mod and activation in bilateral premotor and frontal opercular cortex, and negative correlations were found in left medial prefrontal regions. In addition, positive correlations between disfluency rate and normal–baseline were found in right perisylvian regions, left putamen, and bilateral ventral anterior thalamus (VA)/ventral lateral thalamus and inferior cerebellum.

### Functional Connectivity Analyses

The set of 116 cortical and subcortical ROIs used as seed in the functional connectivity analyses is illustrated in Supplemental Figures S10 and S11. Within the AWS group, two connections were significantly different in the rhythm condition as compared to the normal condition ( $p_{FDR} < .00043$ ), both involving the cerebellum (see Table 4 and Figure 4). The right dentate nucleus showed an increase in connectivity in the rhythm condition, with a cluster covering right cerebellar lobule VI and crus I, as well as vermis VI, while the left cerebellar lobule VIIIa displayed reduced connectivity in the rhythm condition, with a cluster in the left anterior middle frontal gyrus. To determine whether these differences were specific to AWS, a post hoc analysis found that these connections did not reach significance in the ANS group, even using an uncorrected  $\alpha$  level of .05. Instead, ANS had different connections that were significantly

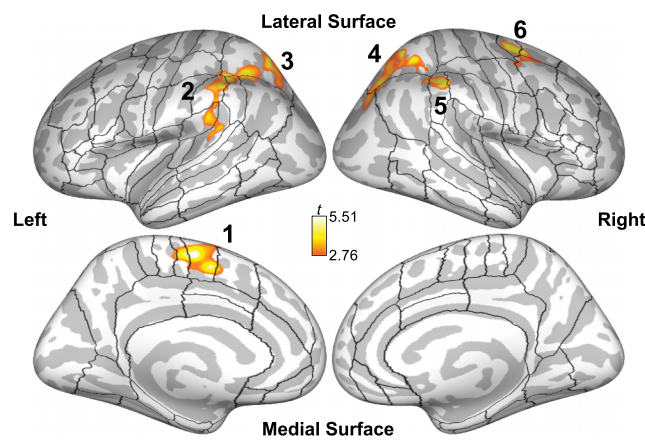
**Table 3.** Cortical clusters with activation differences between the rhythm and normal conditions collapsed across groups (vertex-wise  $p < .01$ , cluster-wise  $p_{FDR} < .05$ ).

Cluster	Peak MNI coordinates			Cluster mass	$p_{FDR}$
	x	y	z		
<i>Combined groups, rhythm &gt; normal</i>					
L lateral superior parietal cortex (aSMg, SPL, PO, PT, pSMg)	–42	–39	45	42456	.0080
R superior parietal cortex (SPL, OC)	32	–51	55	29904	.0090
L supplementary motor area (SMA, dMC, pre-SMA)	–09	–08	59	19658	.0166
L posterior superior parietal cortex (SPL)	–21	–68	59	13488	.0253
R posterior supramarginal gyrus (pSMg, AG)	48	–32	46	12479	.0253
R dorsal premotor cortex (mdPMC, adPMC, pMFg)	23	–04	53	12169	.0253

Note. FDR = false discovery rate; MNI = Montréal Neurological Institute; L = left; aSMg = anterior supramarginal gyrus; SPL = superior parietal lobule; PO = parietal operculum; PT = planum temporale; pSMg = posterior supramarginal gyrus; R = right; OC = occipital cortex; SMA = supplementary motor area; dMC = dorsal primary motor cortex; pre-SMA = presupplementary motor area; AG = angular gyrus; mdPMC = middle dorsal premotor cortex; adPMC = anterior dorsal premotor cortex; pMFg = posterior middle frontal gyrus.



**Figure 3.** Cortical clusters significantly more active during the rhythm condition than the normal condition collapsed across both groups and displayed on an inflated cortical surface (vertex-wise  $p < .01$ , cluster-wise  $p_{FDR} < .05$ ). 1 = left supplementary motor area; 2 = left lateral superior parietal cortex; 3 = posterior superior parietal cortex; 4 = right superior parietal cortex; 5 = right posterior supramarginal gyrus; 6 = right dorsal premotor cortex. Black outlines indicate cortical regions of interest used in the exploratory analysis. FDR = false discovery rate.



different between conditions. Increased connectivity was found in the rhythm condition between the right putamen and a cluster in the anterior cingulate gyrus straddling the midline, right anterior insula and a cluster in the left inferior frontal sulcus, and between the left Heschl's gyrus, right pre-SMA, and right ventral somatosensory cortex seeds and clusters in the left posterior SPL abutting occipital cortex. There was also decrease in connectivity during the rhythm condition between the left inferior temporo-occipital cortex (ITO) and left IFG pars opercularis and triangularis, and between the right anterior dPMC and bilateral occipital cortex (see Supplemental Figure S13).

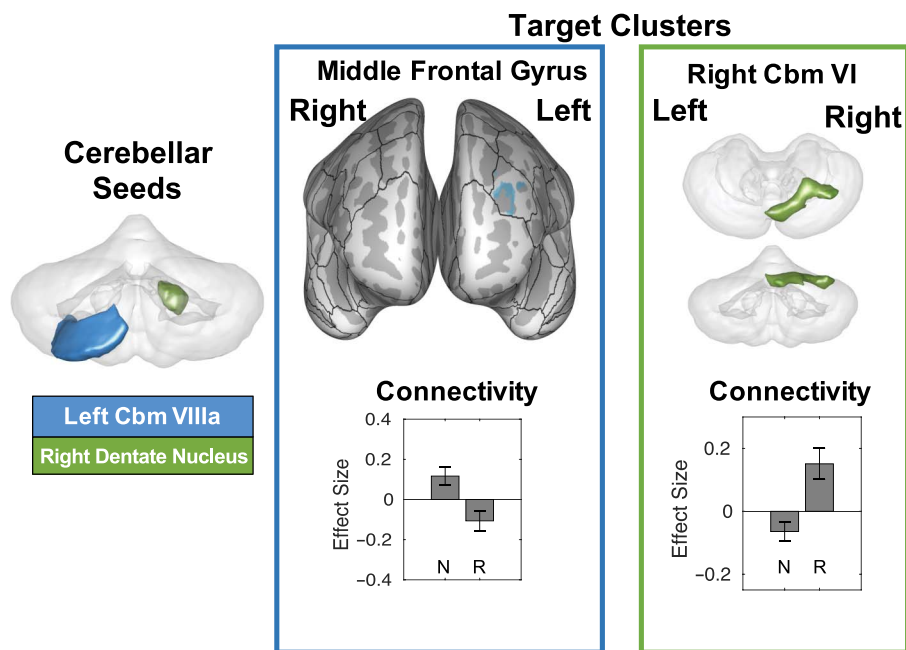
There were three connections that showed a significant interaction between group and speech conditions (normal and rhythm; see Supplemental Figure S12). Connections that were lower in the rhythm condition for AWS and greater in this condition for ANS included the left cerebellar lobules I–IV to the left medial rolandic cortex and precuneus (see result cluster labeled 1 in the bottom-left panel of Supplemental Figure S12) and left VA to right lingual gyrus and occipital cortex (extending to right cerebellar lobule VI; Cluster 2). A connection that was greater in the rhythm condition for AWS and lesser in this condition for ANS was between the right anterior inferior temporal gyrus and a cluster covering parts of the left central operculum, insula, and surrounding

**Table 4.** Functional connectivity analysis—condition and interaction effects.

Seed ROI	Target cluster regions	Peak MNI coordinates			Cluster size (no. of voxels)	$p_{FDR}$
		x	y	z		
AWS, rhythm > normal						
R dentate nucleus	Superior cerebellum (R VI, Ver VI, R crus I)	14	-72	-20	435	$< 1 \times 10^{-6}$
AWS, normal > rhythm						
L Cbm VIIIa	Left anterior middle frontal gyrus (L aMFg)	-28	34	30	170	.000207
ANS, rhythm > normal						
L H	Left parieto-occipital cortex (L SPL)	-24	-70	38	186	.000195
R vSC	Left parieto-occipital cortex (L SPL, L AG, L OC)	-24	-66	34	199	.000176
R aINS	Left inferior frontal sulcus (L aIFs, L pIFs, L aMFg)	-50	26	26	243	.000032
R putamen	Midline cingulate motor cortex (L dCMA, R vCMA, L vCMA, R dCMA, L SMg, L aCG, R SFg, R aCG)	02	14	30	353	$< 1 \times 10^{-6}$
R pre-SMA	Left parieto-occipital cortex (L SPL, L OC, L PCN, L AG)	-16	-70	34	224	.000100
ANS, normal > rhythm						
L ITO	Left inferior frontal gyrus (L vIFo, L pIFt, L pFO)	-56	24	04	175	.000221
Group × Condition interaction						
L Cbm I–IV	Medial sensorimotor cortex (L dSC, L pCG, L PCN, L dMC)	-16	-52	44	254	.000301
L VA	Right inferior occipital cortex (R OC, R LG, R TOF, R VI, Ver VI)	08	-72	-08	365	.000023
R aITg	Left frontoparietal operculum (L aINS, L aCO, L pCO, L pINS, L pFO, L H, L PO, L vPMC)	-40	-12	14	507	$< 1 \times 10^{-6}$

*Note.* Roman numerals indicate cerebellar lobules. ROI = region of interest; MNI = Montréal Neurological Institute; FDR = false discovery rate; AWS = adults who stutter; R = right; Ver = vermis; L = left; Cbm = cerebellum; aMFg = anterior middle frontal gyrus; ANS = adults who do not stutter; H = Heschl's gyrus; SPL = superior parietal lobule; vSC = ventral primary somatosensory cortex; AG = angular gyrus; OC = occipital cortex; aINS = anterior insula; aIFs = anterior inferior frontal sulcus; pIFs = posterior inferior frontal sulcus; dCMA = dorsal cingulate motor area; vCMA = ventral cingulate motor area; SMg = supramarginal gyrus; aCG = anterior cingulate gyrus; SFg = superior frontal gyrus; pre-SMA = presupplementary motor area; PCN = precuneus; ITO = inferior temporo-occipital cortex; vIFo = ventral inferior frontal gyrus pars opercularis; pIFt = posterior inferior frontal gyrus pars triangularis; pFO = posterior frontal operculum; dSC = dorsal primary somatosensory cortex; pCG = posterior cingulate gyrus; dMC = dorsal primary motor cortex; VA = ventro-anterior regions of the thalamus; LG = lingual gyrus; TOF = temporo-occipital fusiform gyrus; aITg = anterior inferior temporal gyrus; aCO = anterior central operculum; pCO = posterior central operculum; pINS = posterior insula; PO = parietal operculum; vPMC = ventral premotor cortex.

**Figure 4.** A summary of functional connections that are significantly different between the normal and rhythm conditions in adults who stutter. Seed regions for these connections are indicated on the left side on a transparent 3D rendering of the cerebellum (viewed posteriorly), and colors in the rest of the figure refer back to these seed regions. Two target clusters (representing two distinct connections) are displayed in the right portion of the figure. Target Cluster 1 is projected onto an inflated surface of cerebral cortex (anterior view), along with the full cortical region of interest parcellation of the SpeechLabel atlas described in Cai et al. (2014). Target Cluster 2 is displayed on a transparent 3D rendering of the cerebellum (top view: superior; bottom view: posterior). The connectivity effect sizes in the normal and rhythm conditions for each connection are displayed below each cluster visualization. Roman numerals indicate cerebellar lobules. Error bars indicate 90% confidence intervals. Cbm = cerebellum; N = normal; R = rhythm.



regions. Simple effects from each group and condition are shown in the bottom panel of Supplemental Figure S12. On the basis of the results that showed increased connectivity for AWS between different parts of the cerebellum during isochronous speech, we performed a test comparing average pairwise connectivity among all 20 cerebellar ROIs active during speech. This test revealed that these ROIs show a significant Group  $\times$  Condition interaction ( $t = 2.73, p = .011$ ), driven by an increase in connectivity for AWS from normal to rhythm ( $t = 2.68, p = .019$ ) and a nonsignificant decrease in connectivity for ANS ( $t = -1.93, p = .073$ ).

For the AWS group, there were multiple functional connections that were significantly correlated with either SSI-Mod or disfluency rate. Results are summarized in Table 5 and Supplemental Figures S14–S19. Of note, connectivity differences between the normal and baseline conditions were negatively correlated with SSI-Mod between cerebellar vermis crus II and bilateral cerebellum lobules IX and VIIIb (see Figure 5). There was also a significant positive correlation between SSI-Mod and connections between the left cerebellar lobule VIIIa and right IFG pars orbitalis (IFr) in the normal condition compared to baseline (see Figure 5). In addition, connectivity differences between the rhythm and baseline conditions were negatively correlated with SSI-Mod between the right temporoparietal junction and cluster in each of the left anterior SMg (aSMg), left IFr,

left ITO, and right VA (see Figure 6), and between the right PT and a cluster in medial premotor cortex/SMA (see Figure 7).

## Discussion

This study aimed to characterize the changes in functional activation and connectivity that occur when adults time their speech to an external metronomic beat and how these changes differ in AWS compared to ANS. Extending previous work, this paradigm was novel in that the metronome was paced at the typical rate of English speech. The rate and isochronicity of paced speech by AWS was also similar to that of ANS. Consistent with prior literature, AWS produced significantly fewer disfluencies during externally paced speech than during normal, internally paced speech (see Figure 2). Controlling for speaking rate, participants exhibited greater activation during isochronously paced speech than internally paced speech in left hemisphere sensory association areas and bilateral attentional and premotor regions. AWS had greater functional connectivity during isochronous speech than internally paced speech within the cerebellum and reduced connectivity between the left inferior cerebellum and left prefrontal cortex. Finally, there were significant correlations between SSI-Mod and functional connections within the cerebellum, between the cerebellum and

**Table 5.** Functional connectivity analysis—correlations with SSI-Mod (a modified version of the Stuttering Severity Index–Fourth Edition) and disfluency rate.

Seed ROI	Target cluster regions	Peak MNI coordinates			Cluster size (no. of voxels)	<i>P</i> <sub>FDR</sub>
		<i>x</i>	<i>y</i>	<i>z</i>		
AWS, normal > baseline, negative correlation with SSI-Mod						
L aSMg	Left parietal operculum (L PO, L aSMg)	-32	-22	24	216	.000108
L aCO	Left medial prefrontal cortex (L SFg, L aCG)	-12	38	30	227	.000031
L OC	Right medial posterior temporal lobe (R LG, R TOF, R pPHg, R pTFg)	22	-54	-12	320	.000001
L Cbm crus I	Right inferior temporal lobe (R LG, R TOF, R pPHg, R pTFg, R ITO, R avSTs, R adSTs, R VI, R V, R MGN, R SN, R VPM)	38	-42	-14	1672	< 1 × 10 <sup>-6</sup>
	Left inferior medial temporal lobe (L LG, L TOF, L pTFg, L V, L VI, L I-IV, R I-IV)	-06	-52	-10	396	< 1 × 10 <sup>-6</sup>
L STh	Left occipital cortex/superior cerebellum (L TOF, L OC, L VI, L crus I, Ver VI)	-8	-78	-18	235	.000178
Cbm vermis crus II	Brainstem/inferior cerebellum (brainstem, L IX, R VIIIb, R IX)	-4	-52	-60	174	.000364
R midMC	Left medial prefrontal cortex (L SFg, L FP)	02	54	-08	231	.000109
R VA	Left angular gyrus (L AG)	-58	-64	28	264	.000026
	Right middle temporal–occipital cortex (R MTO)	60	-60	02	253	.000026
	Right parietal operculum (R PO, R PT, R pINS)	34	-24	16	240	.000029
R vSC	Left medial prefrontal cortex (L SFg, L FP)	-04	52	18	353	.000001
R pdSTs	Midline medial prefrontal cortex (L SFg, L FP, R SFg)	-12	40	26	184	.000247
AWS, normal > baseline, positive correlation with SSI-Mod						
L Cbm Villa	Right orbital inferior frontal gyrus (R IFr, R FP, R FOC)	50	50	-12	234	.000061
L plFt	Right angular gyrus (R AG)	36	-70	46	271	.000017
R VA	Right inferior temporal sulcus (R pMTg, R plTg)	54	-38	-16	169	.000391
R vSC	Brainstem (brainstem)	08	-14	-36	292	.000005
	Right posterior angular gyrus (R AG, R OC)	36	-74	36	260	.000010
AWS, rhythm > baseline, negative correlation with SSI-Mod						
L aSMg	Right parietal operculum (R PO, R aSMg, R pINS)	38	-24	24	404	< 1 × 10 <sup>-6</sup>
L FOC	Right inferior parietal cortex (R aSMg, R PO, R vSC)	50	-22	40	374	< 1 × 10 <sup>-6</sup>
L ITO	Right dorsal rolandic cortex (R dMC, R dPMC, R vSC)	18	-10	70	358	< 1 × 10 <sup>-6</sup>
	Right frontal operculum (R pFO, R aFO, R aINS)	32	14	10	299	.000002
	Right temporo-parietal junction (R PT, R PO)	40	-30	30	290	.000002
L Cbm I-IV	Midline occipital cortex (R OC, L OC, R LG)	00	-70	06	401	< 1 × 10 <sup>-6</sup>
L Cbm crus I	Left parieto-occipital fissure (L PCN, L OC)	-08	-64	28	259	.000073
R TP	Right middle frontal gyrus (R pMFg, R plFS)	52	18	38	213	.000254
R PT	Midline medial precentral gyrus (R dPMC, L dPMC, L dMC, L SMA)	10	-28	74	190	.000194
R aCG	Right supramarginal gyrus (R pSMg, R AG)	54	-40	46	350	.000001
R midMC	Midline rostral prefrontal cortex (L aCG, L FP, R FP, R FMC)	-08	50	04	370	< 1 × 10 <sup>-6</sup>
R STh	Right anterior insula (R aINS, R IFr)	34	22	12	368	.000001
R SN	Right inferior cerebellum (R IX, R X, L IX, Ver VIIIa, Ver VIIIb, L VIIb, R VIIIa, L VIIIa, L dentate, R VIIb, Ver IX)	-10	-66	-40	404	< 1 × 10 <sup>-6</sup>
R VA	Right temporo-parietal junction (R PO, R PT, R aSMg, R vSC)	60	-22	20	559	< 1 × 10 <sup>-6</sup>
	Right middle temporo-occipital cortex (R MTO)	60	-50	02	232	.000036
R Cbm I-IV	Midline occipital cortex (L OC, R OC, R PCN)	-08	-78	06	286	.000008
Cbm vermis VI	Left parieto-occipital fissure (L PCN, L OC, L LG)	-24	-62	24	622	< 1 × 10 <sup>-6</sup>
	Right parieto-occipital fissure (R PCN, R OC)	10	-64	16	382	< 1 × 10 <sup>-6</sup>
	Right anterior cingulate cortex (R aCG)	16	26	20	160	.000415
AWS, rhythm > baseline, positive correlation with SSI-Mod						
L aSMg	Left inferior occipital cortex (L OC, L TOF, L ITO)	-42	-74	-16	226	.000032

(table continues)

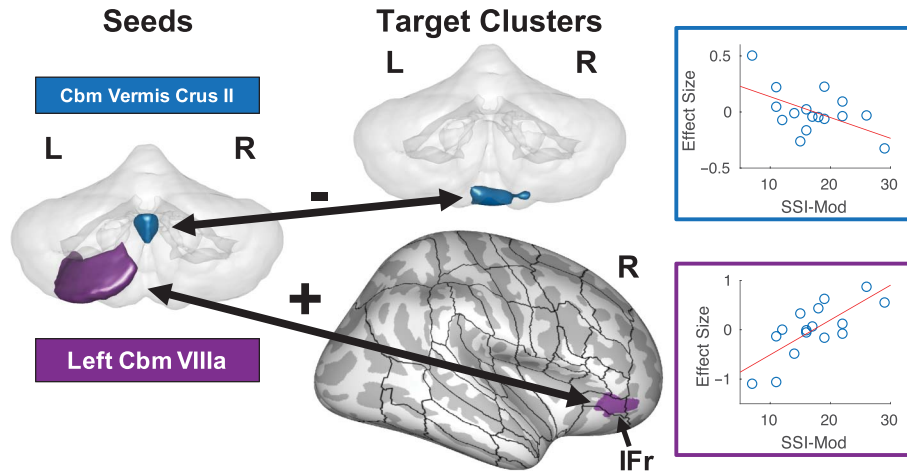
Table 5. (Continued).

Seed ROI	Target cluster regions	Peak MNI coordinates			Cluster size (no. of voxels)	$p_{FDR}$
		x	y	z		
AWS, normal > baseline, negative correlation with disfluency rate						
L pdPMC	Right insula (R alNS, R aCO, R putamen)	42	04	08	186	.000360
L pdSTs	Right medial temporal cortex (R pPHg, R pTFg)	34	-34	-20	191	.000059
	Right lateral occipital cortex (R OC)	38	-64	-12	144	.000292
L pCO	Midline rolandic cortex (R dMC, R dPMC, R dSC, L dMC, L dSC)	14	-22	78		
L dentate	Right occipital cortex (R OC)	12	-84	16	216	.000083
L Cbm crus I	Right dorsal prefrontal cortex (R FP, R SFg)	24	48	36	199	.000038
Cbm vermis crus I	Right occipital cortex (R OC, R PCN)	06	-80	28	1149	$< 1 \times 10^{-6}$
Cbm vermis VIIIa	Right occipital cortex (R OC, R LG)	10	-62	06	184	.000223
AWS, normal > baseline, positive correlation with disfluency rate						
R Cbm X	Inferior temporal cortex (R pITg, R pTFg, R VI)	46	-36	-28	356	.000001

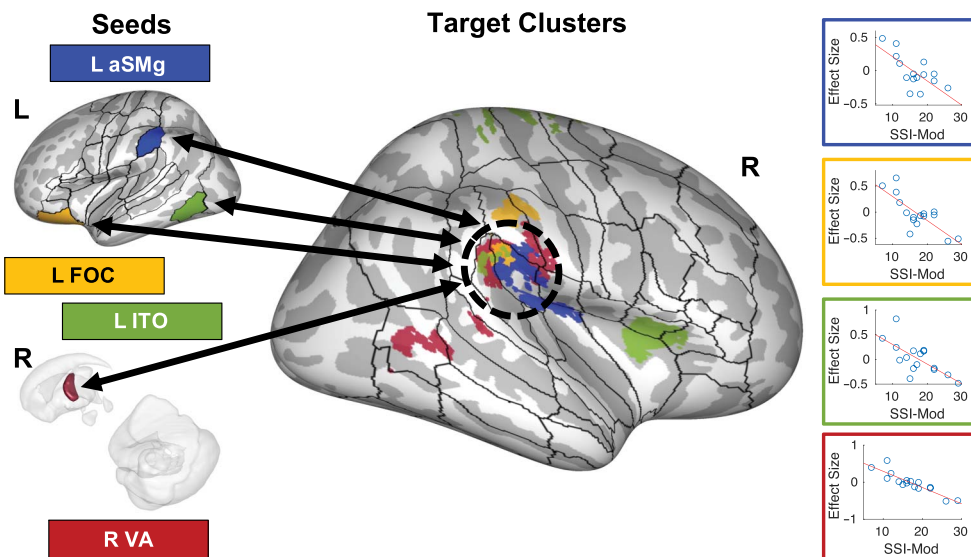
*Note.* Roman numerals indicate cerebellar lobules. Regions of the SpeechLabel atlas (Cai et al., 2014) containing at least 10 voxels of a given cluster are indicated in parentheses. ROI = region of interest; MNI = Montréal Neurological Institute; FDR = false discovery rate; L = left; aSMg = anterior supramarginal gyrus; PO = parietal operculum; aCO = anterior central operculum; SFg = superior frontal gyrus; aCG = anterior cingulate gyrus; OC = occipital cortex; R = right; LG = lingual gyrus; TOF = temporo-occipital fusiform gyrus; pPHg = posterior parahippocampal gyrus; pTFg = posterior temporal fusiform gyrus; Cbm = cerebellum; ITO = inferior temporo-occipital cortex; avSTs = anterior ventral superior temporal sulcus; adSTs = anterior dorsal superior temporal sulcus; MGN = medial geniculate nucleus of the thalamus; SN = substantia nigra; VPM = ventral postero-medial portion of the thalamus; STh = subthalamic nucleus; Ver = vermis; midMC = middle primary motor cortex; FP = frontal pole; VA = ventral anterior portion of the thalamus; AG = angular gyrus; MTO = middle temporo-occipital cortex; PT = planum temporale; pINS = posterior insula; vSC = ventral primary somatosensory cortex; pdSTs = posterior dorsal superior temporal sulcus; IFR = inferior frontal gyrus pars orbitalis; FOC = fronto-orbital cortex; pIFt = posterior inferior frontal gyrus pars triangularis; pMTg = posterior middle temporal gyrus; pITg = posterior inferior temporal gyrus; dMC = dorsal primary motor cortex; dPMC = dorsal premotor cortex; pFO = posterior frontal operculum; aFO = anterior frontal operculum; alNS = anterior insula; PCN = precuneus; TP = temporal pole; pMFg = posterior middle frontal gyrus; pIFs = posterior inferior frontal sulcus; SMA = supplementary motor area; pSMg = posterior supramarginal gyrus; FMC = fronto-medial cortex; pdPMC = posterior dorsal premotor cortex; dSC = dorsal primary somatosensory cortex.



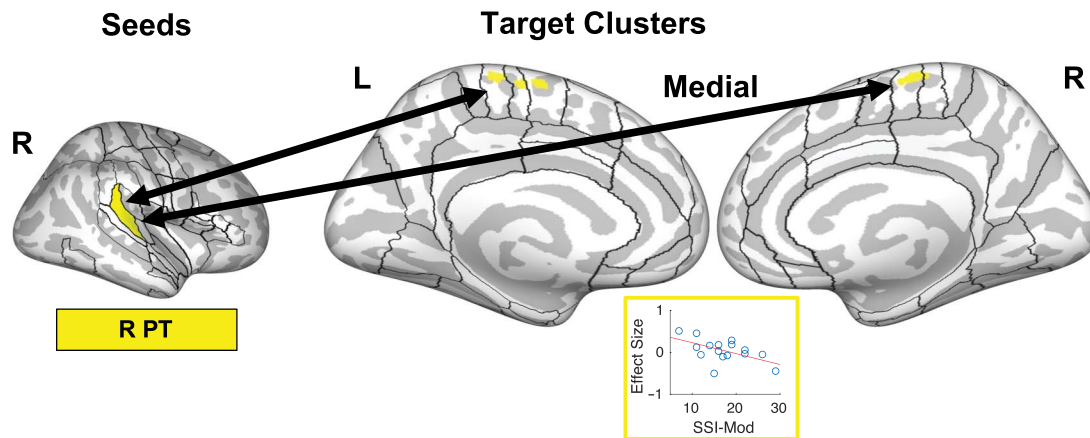
**Figure 5.** Two notable correlations of cerebellar functional connectivity (normal > baseline) with stuttering severity. Seed regions for these connections are indicated on the left side of the figure on a transparent 3D rendering of the cerebellum viewed posteriorly. Colors in the rest of the figure refer back to these seed regions. Target clusters are either displayed on the same transparent rendering of the cerebellum or projected onto an inflated surface of cerebral cortex, along with the full cortical region of interest parcellation of the SpeechLabel atlas described in Cai et al. (2014). The “+” and “-” indicate positive and negative correlations, respectively. The right portion of the figure plots the beta estimates of the psychophysiological interaction regressors from individual adults who stutter against stuttering severity. Roman numerals indicate cerebellar lobules. Full results of this analysis can be found in Supplemental Figures S14 and S15 and Table 5. Cbm = cerebellum; IFr = inferior frontal gyrus pars orbitalis; L = left; R = right; SSI-Mod = modification of the Stuttering Severity Instrument–Fourth Edition score.



**Figure 6.** Correlations of functional connectivity (rhythm > baseline) between seed regions of interest (ROIs) and the right temporoparietal junction with stuttering severity. Seed regions for these connections are indicated on the left side of the figure either on an inflated surface of the left cerebral cortex or on a transparent 3D rendering of right subcortical structures viewed medially. Colors in the rest of the figure refer back to these seed regions. Target clusters are projected onto an inflated surface of the right cerebral cortex, along with the full cortical ROI parcellation of the SpeechLabel atlas described in Cai et al. (2014). The black dashed oval indicates a rough border of the right temporoparietal junction. The right portion of the figure plots the beta estimates of the psychophysiological interaction regressors from individual adults who stutter against stuttering severity for each functional connection. Full results of this analysis can be found in Supplemental Figures S16 and S17 and Table 5. aSMg = anterior supramarginal gyrus; FOC = fronto-orbital cortex; ITO = inferior temporo-occipital junction; L = left; R = right; SSI-Mod = modification of the Stuttering Severity Instrument–Fourth Edition score; VA = ventral anterior portion of the thalamus.



**Figure 7.** Correlations of functional connectivity (rhythm > baseline) between the right planum temporale and a medial premotor cortex with stuttering severity. The seed region is indicated on the left side of the figure on an inflated surface of the right cerebral cortex. One target cluster (straddling the midline) is projected onto an inflated surface of the cerebral cortex, along with the full cortical region of interest parcellation of the SpeechLabel atlas described in Cai et al. (2014). Below, the beta estimates of the psychophysiological interaction regressors from individual adults who stutter are plotted against stuttering severity. Full results of this analysis can be found in Supplemental Figures S16 and S17 and Table 5. L = left; PT = planum temporale; R = right; SSI-Mod = modification of the Stuttering Severity Instrument–Fourth Edition score.



orbitofrontal cortex, and between the right temporoparietal junction and left hemisphere speech-related regions. The following sections discuss these results in relation to prior behavioral and neuroimaging literature.

### ***A Possible Compensatory Role for the Cerebellum in AWS***

A role for the cerebellum in mediating speech timing is well established (see Ackermann, 2008, for a review), and damage to this structure can lead to “scanning speech,” where syllables are evenly paced (Duffy, 2013). Previous work posits that when the basal ganglia–SMA “internal” timing system is impaired in AWS, the cerebellum, along with lateral cortical premotor structures, forms part of an “external” timing system that is recruited (Alm, 2004; Etchell et al., 2014). In support of this, numerous fMRI and positron emission tomography studies demonstrate cerebellar overactivation and hyperconnectivity during normal speech production in AWS (e.g., Brown et al., 2005; Chang et al., 2009; Ingham et al., 2012; Lu et al., 2012; Lu, Peng, et al., 2010; Watkins et al., 2007) that is reduced following therapy (De Nil et al., 2001; Lu et al., 2012; Neumann et al., 2003; Toyomura et al., 2015), a potential indication of an organic attempt at compensation. In this study, the increased connectivity among speech-related regions of the cerebellum, along with increased fluency during the rhythm condition, may thus reflect similar neural processes.

It should be noted that this functional connectivity does not reflect direct structural connectivity between a seed and target region. As suggested by Bernard et al. (2013), we interpret the result of increased within-cerebellar connectivity as reflecting an increase in synchrony among multiple cerebrocerebellar loops. Thus, in AWS, areas of cerebral cortex

may simultaneously impinge on distinct areas of cerebellum to utilize the cerebellum’s temporal processing capabilities to ensure accurate speech timing during the rhythm condition.

The reduction in connectivity between the left prefrontal cortex and inferior cerebellum connection may be an exception. Both regions are functionally connected during rest with areas of the ventral attention network, including bilateral temporoparietal junction and IFG (Buckner et al., 2011; Vossel et al., 2014; Yeo et al., 2011). This network is associated with modulating attention based on new or surprising stimuli (Vossel et al., 2014), is largely right-lateralized (Vossel et al., 2014), and overlaps with regions involved in responding to sensory feedback errors during speech production (Golfinopoulos et al., 2011; Tourville et al., 2008). Indeed, cerebellar lobule VIII is also involved in sensory feedback control (Golfinopoulos et al., 2011; Tourville et al., 2008) and suprasyllabic speech sequencing (Bohland & Guenther, 2006). Thus, a reduction in connectivity between these two regions during the rhythm condition may reflect a decrease in reliance on this network in favor of more top-down control in AWS.

### ***Changes in Activation During Isochronous Speech***

Comparing neural activation between isochronously paced and normal speech showed that subjects had greater activation during isochronous speech in left hemisphere medial premotor and sensory association areas, bilateral parietal cortex, and right hemisphere dPMC. Activation in the left temporoparietal sensory association cortex (PT, aSMG) and right vPMC (in the exploratory results) may be related to increased reliance on sensory feedback control during this novel speech condition. Previous studies have shown that sensory feedback errors (i.e., mismatches between the auditory

signal expected from the current motor commands and the actual auditory signal) lead to increased activation in secondary auditory and somatosensory areas (Hashimoto & Sakai, 2003; Parkinson et al., 2012; Takaso et al., 2010; Tourville et al., 2008), whereas greater activation in right vPMC is thought to reflect the transformation of sensory errors into corrective motor responses (Golfinopoulos et al., 2011; Hashimoto & Sakai, 2003; Tourville et al., 2008). Temporoparietal cortex may also play a more general role in audio–motor integration (Hickok et al., 2003); therefore, increased activation in this region may be indicative of the need to hold the rhythmic auditory stimulus in working memory and translate it into a motoric response in the rhythm condition of the current study. This is supported by increased activity in bilateral intraparietal sulcus and posterior SMg, additional regions commonly recruited in working memory tasks (Rottschy et al., 2012).

There was also increased activation during isochronous speech in areas thought to be involved in speech planning and sequencing (left SMA; Bohland et al., 2010; Civier et al., 2013; Guenther, 2016), producing complex motor sequences (left SPL; Haslinger et al., 2002; Heim et al., 2012), producing novel sequences (left SPL; Jenkins et al., 1994; Segawa et al., 2015), attending to stimulus timing (left SPL; Coull, 2004), and controlled respiration (right dPMC; McKay et al., 2003). The rhythm condition requires participants to produce speech in an unfamiliar way. This change in their speech production results in speech becoming less automatic and may require greater recruitment in these areas for timing the sequence of syllables (Alario et al., 2006; Bohland & Guenther, 2006; Schubotz & von Cramon, 2001). Bengtsson et al. (2004, 2005) found that, for both finger tapping and simple repetition of “pa,” more complex timing led to increased activation in SMA compared to simple patterns. The increased need to implement a timing pattern recruited the same structure that mediates temporal sequencing.

Unlike previous studies (Braun et al., 1997; Stager et al., 2003; Toyomura et al., 2011, 2015), AWS did not exhibit significantly increased activation in the rhythm condition compared to the normal condition. The most consistent finding from these studies was that both groups showed increased activation in bilateral auditory regions during isochronously paced speech and that AWS showed greater increases in the basal ganglia. In this study, the lack of clear between-conditions effects within the AWS group or between the AWS and ANS groups may be due to more individual variability for AWS than ANS for this contrast. Future work is needed to determine whether this within-group variability is driving the null findings in the AWS group. Furthermore, Toyomura et al. (2011) found that, while areas of the basal ganglia, left precentral gyrus, left SMA, left IFG, and left insula were less active in AWS during normal speech, activity in these areas increased to the level of ANS during isochronous speech. These results suggested that isochronously paced speech had a “normalizing” effect on activity in these regions, which differs with the present results.

There are methodological differences between the current work and similar studies that also could have impacted

the results. In the current study, the rhythmic stimulus was presented prior to speaking regardless of the condition, unlike previous work in which the participant heard the stimulus while speaking and only during the rhythm condition (Toyomura et al., 2011). Thus, group effects reported by Toyomura et al. (2011) likely reflect auditory processing of the pacing stimulus in addition to any differences in speech motor processes. Second, our study sought to examine the rhythm effect when speech was produced at a conversational speaking rate. Previous studies used a metronome set at 92–100 beats per minute, considerably slower than the mean conversational rate in English (228–372 syllables per minute; Davidow, 2014; Pellegrino et al., 2011) and the rate observed in our study (approximately 207 syllables per minute). While Toyomura et al. (2011, 2015) instructed participants to speak at a similar rate during the normal condition (when previous studies had not), the slower tempo overall may have led to increased auditory feedback processing. This could have modified the mechanisms by which ANS and AWS controlled their speech timing. Finally, only one of the previous studies accounted for disfluencies during the task in their imaging analysis (Stager et al., 2003), despite significant correlations with brain activation (Braun et al., 1997). However, given the small number of disfluencies in this and previous studies, this effect may have had a limited impact on the results.

### *Correlation Between Activation and Severity*

The voxel/vertex-based analysis did not find significant correlations between disfluency rate and activation in the normal–baseline contrast. However, in the exploratory ROI analysis, activation in the left VA thalamus and bilateral ventral lateral thalamus had among the strongest positive correlations with the disfluency rate ( $p < .005$ ; see Supplemental Table S4 and Supplemental Figure S9 for details). These nuclei are part of both the cortico-cerebellar and cortico-basal ganglia motor loops and are structurally connected with premotor and primary motor areas (Barbas et al., 2013). As relays between subcortical structures and the cortex, increased activation for participants with a higher disfluency rate during the task may reflect greater reliance upon these modulatory pathways during speech. It is also worth noting that, with an exploratory threshold ( $p < .05$ , uncorrected), some ROIs follow similar patterns to previous literature. The propensity to stutter during the task, measured by the disfluency rate, was associated with greater cortical activation in largely right hemisphere regions and bilateral subcortical activation at uncorrected thresholds. The right-lateralized cortical associations in this study may reflect increased compensatory activity in AWS (as in Braun et al., 1997; Cai et al., 2014; Kell et al., 2009; Preibisch et al., 2003; Salmelin et al., 2000). This is supported by the fact that fluency-inducing therapy leads to more left-lateralized activation (De Nil et al., 2003; Neumann et al., 2003, 2005), similar to that of neurotypical speakers. It should be noted that, due to the low number of disfluencies exhibited during the task, determining a clear relationship between fluency and activation may not have been possible.

Functional connectivity between multiple seed ROIs and target clusters was significantly correlated with SSI-Mod. Given the large number of these significantly correlated connections, we focus here on what we consider to be the most salient findings; further detail regarding the full set of findings is provided in the supplemental materials.

When comparing the normal and baseline conditions, the negative association between SSI-Mod and the connection between cerebellar vermis crus II and midline inferior cerebellum indicates that less severe AWS have greater within-cerebellum connectivity. This fits conceptually with the result of increased connectivity within the cerebellum during the rhythm condition—both conditions associate the cerebellum with greater fluency. There was also a positive correlation between SSI-Mod and the connection between the left cerebellum VIIIa and right IFR. The direction of this connection is surprising given previous work. For instance, Sitek et al. (2016) found a negative relationship between SSI scores and the connection between the left cerebellum and IFR in resting-state connectivity, and in Kell et al. (2018), there was hyperconnectivity between the cerebellum and left IFR in the comparison between overt and covert speech for recovered AWS. These suggest that greater fluency was associated with enhanced connections between these regions. However, the cerebellar regions involved in these connections were not as fine-grained as the ROI in the current study, and the specific tasks on which these connections were based were different than the normal–baseline comparison in this study.

During the rhythm condition compared to baseline, multiple connections—between the left fronto-orbital cortex, left aSMg, left ITO, and right VA seeds and overlapping clusters in the right temporoparietal junction—were negatively correlated with SSI-Mod. Thus, more severe AWS had lower connection strengths compared to less severe AWS. In general, these connections support the idea that the right hemisphere is recruited to compensate for impaired left hemisphere processing (Braun et al., 1997; De Nil et al., 2000; Fox et al., 1996). Indeed, this temporoparietal region was found to be hyperactive in a meta-analysis of stuttering neuroimaging studies (Belyk et al., 2015). The convergence of these connections, specifically in the right temporoparietal junction, may imply association with this region's role in responding to salient or unexpected events (Corbetta & Shulman, 2002). In the realm of speech production, these connections (especially with left aSMg) may reflect increased use of the somatosensory feedback loop by less severe AWS to control speech during the rhythm condition (Golfinopoulos et al., 2011). One additional negative correlation worth mentioning for the rhythm–baseline contrast is the negative association between SSI-Mod and the connection between the right PT and midline primary motor cortex/premotor cortex/SMA. In the auditory feedback loop, as proposed by Tourville et al. (2008) and Guenther (2016), sensory state, target, and error maps send error signals to the right PMC to generate connective motor commands. Connectivity between the right PT and medial premotor regions may then reflect an interface between these sensory feedback loops

and the SMA–basal ganglia “internal” timing system, which is disrupted in stuttering (Chang & Guenther, 2020). More fluent speakers may use this connection to a greater extent in order to resolve conflicts between competing motor programs (Guenther, 2016).

### *Limitations*

One potential limitation to this study was that trial types were pseudorandomly presented within a given run. Since the sequence of tones was presented before every trial and the participants did not know the condition ahead of time, participants needed to refrain from speaking at the pace of the tone sequence during normal trials. This process of ignoring the tone sequence during production of their sentence may have recruited additional brain areas for the normal condition only, potentially confounding the neural response. However, presenting the tone sequence before every trial was done specifically to eliminate the confound of tone sequence auditory processing found in previous studies. Even if the rhythm trials and normal trials were presented in a blocked fashion, such that participants knew the condition ahead of time, they would still have to either ignore the tones on the normal trials or risk the confound of attending to the tones in one condition and not the other. As it is, there are a few indications that this contrast reflects the difference in speaking styles between conditions. First, the reduction in disfluencies in the “rhythm” condition compared to the “normal” condition shows that the fluency-enhancing effect took place. Thus, any neural changes between the conditions could plausibly reflect this effect. Second, the pattern of pacing tones that participants hear is quite simple and is the same throughout the experiment. Furthermore, the task is well practiced by the participants from a similar behavioral experiment that they participated in prior to the fMRI task. Thus, listening to the tones before each trial is merely a reminder of the pace rather than something that requires significant attentional resources. Finally, all significant corrected results and most exploratory results from the rhythm–normal activation contrast demonstrated greater activation in the rhythm condition. If there were additional areas recruited for ignoring a rhythm when producing speech, they would have probably led to greater activation during the normal condition. That being said, if a given region mediated both isochronous speech production and ignoring an external pacing stimulus, the direction of activation change between conditions would be mixed, which could potentially lead to false negative findings. It is also possible that the reduced connectivity between the left anterior prefrontal cortex and left inferior cerebellum in the rhythm may reflect this additional “ignoring” process during the normal condition. Balancing the need to avoid the confounds of the auditory stimulus presentation and the process of ignoring the tones in the unpaced condition is a challenge that will need to be addressed in future work.

In addition, as mentioned in the Data Acquisition section, while the sparse sampling paradigm allows participants to hear themselves speak without addition scanner



noise and decouples the functional acquisition from task-related motion, collecting a single data point per trial poses some challenges to interpretation of the results. One challenge is the assumption that the single acquisition captures the peak of the BOLD response, which has been shown to vary across brain regions and participants (Handwerker et al., 2004; Janssen & Mendieta, 2020). This is an issue common to many sparse sampling paradigms and implies that, because the peak response of some brain regions may not be captured in this single acquisition, there is less power to detect significant results in these regions. For this study, because of the prolonged duration of the sentence production (approximately 2 s) and the relatively slow acquisition of 2.47 s, the single acquisition would provide a broad sampling of the hemodynamic response across a range of different delay times. Furthermore, computing functional connectivity from sparsely sampled data has much less power and temporal resolution than for continuous data. This could negatively impact the detectability of significant connections that would otherwise be found with more scans and a greater sampling of time points that include BOLD response peaks from a broader range of regions and participants. Future studies investigating functional connectivity of speaking tasks that rely on auditory processing and speech production could be improved by acquiring more samples (see Perrachione & Ghosh, 2013, for a discussion of these issues for task activation).

Finally, the current results are not consistent with a recent meta-analysis examining activation differences between AWS and ANS (Belyk et al., 2015, 2017), which found that AWS consistently had overactivation in right hemisphere cortical structures and underactivation in left hemisphere structures, especially in motor and premotor areas. However, this study's exploratory analysis suggested that AWS had decreased activation in the left frontal operculum during the rhythm condition as compared to the ANS group. Previous work has shown gray matter and white matter anomalies in and near the left IFG (Beal et al., 2013, 2015; Chang et al., 2008, 2011; Kell et al., 2009; Lu et al., 2012), which may be related to this underactivation. Based on the exploratory nature of these findings, future work and meta-analytic testing are needed to determine whether these are true population differences.

## Conclusions

In this study, we examined brain activation patterns that co-occur with the introduction of an external pacing stimulus. We found that AWS showed an overall decrease in disfluencies during this condition, as well as functional connectivity changes both within the cerebellum and between the cerebellum and prefrontal cortex. Involvement of these structures suggests that isochronously paced speech activates compensatory timing systems and potentially modulates feedback control and attentional systems. This study provides greater insight into the network of brain areas that support (or respond to) fluency in relation to the rhythm effect and its correspondence to longer term fluency provided through

natural compensation. It is our hope that, in conjunction with the large body of work already published on fluency-enhancing techniques and future studies with more focused analyses, the field will come to a better understanding of the pathophysiology of stuttering and fluency and that this information will be used to provide more targeted treatments and, ultimately, improve quality of life for those who stutter.

## Acknowledgments

The research reported here was supported by National Institute on Deafness and Other Communication Disorders Grants R01DC007683 (F. H. Guenther, Principal Investigator [PI]) and T32DC013017 (training for S. A. Frankford and E. S. Heller Murray; Christopher Moore, PI) and by National Science Foundation Grant NSF1625552 (Boston University Cognitive Neuroimaging Center Research Instrumentation Grant; Chantal Stern, PI). We are grateful to Diane Constantino, Barbara Holland, Matthias Heyne, Megan Thompson, Elaine Kearney, Julianne Leber, and Erin Archibald for their assistance with subject recruitment and data collection and to Ina Jessen, Mona Tong, and Brittany Steinfeld for their help with behavioral data analysis. This work benefited from helpful discussions with, or comments from, other members of the Boston University Speech Lab.

## References

- Ackermann, H. (2008). Cerebellar contributions to speech production and speech perception: Psycholinguistic and neurobiological perspectives. *Trends in Neurosciences*, *31*(6), 265–272. <https://doi.org/10.1016/j.tins.2008.02.011>
- Alario, F.-X., Chainay, H., Lehericy, S., & Cohen, L. (2006). The role of the supplementary motor area (SMA) in word production. *Brain Research*, *1076*(1), 129–143. <https://doi.org/10.1016/j.brainres.2005.11.104>
- Alm, P. A. (2004). Stuttering and the basal ganglia circuits: A critical review of possible relations. *Journal of Communication Disorders*, *37*(4), 325–369. <https://doi.org/10.1016/j.jcomdis.2004.03.001>
- Andersson, J. L. R., Hutton, C., Ashburner, J., Turner, R., & Friston, K. (2001). Modeling geometric deformations in EPI time series. *NeuroImage*, *13*(5), 903–919. <https://doi.org/10.1006/nimg.2001.0746>
- Andrews, G., Howie, P. M., Dozsa, M., & Guitar, B. E. (1982). Stuttering: Speech pattern characteristics under fluency-inducing conditions. *Journal of Speech and Hearing Research*, *25*(2), 208–216. <https://doi.org/10.1044/jshr.2502.208>
- Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *NeuroImage*, *26*(3), 839–851. <https://doi.org/10.1016/j.neuroimage.2005.02.018>
- Azrin, N., Jones, R. J., & Flye, B. (1968). A synchronization effect and its application to stuttering by a portable apparatus. *Journal of Applied Behavior Analysis*, *1*(4), 283–295. <https://doi.org/10.1901/jaba.1968.1-283>
- Barbas, H., García-Cabezas, M. Á., & Zikopoulos, B. (2013). Frontal-thalamic circuits associated with language. *Brain and Language*, *126*(1), 49–61. <https://doi.org/10.1016/j.bandl.2012.10.001>
- Barber, V. (1940). Studies in the psychology of stuttering, XVI: Rhythm as a distraction in stuttering. *Journal of Speech Disorders*, *5*(1), 29–42. <https://doi.org/10.1044/jshd.0501.29>
- Beal, D. S., Gracco, V. L., Brettschneider, J., Kroll, R. M., & De Nil, L. F. (2013). A voxel-based morphometry (VBM) analysis of regional grey and white matter volume abnormalities

- within the speech production network of children who stutter. *Cortex*, 49(8), 2151–2161. <https://doi.org/10.1016/j.cortex.2012.08.013>
- Beal, D. S., Lerch, J. P., Cameron, B., Henderson, R., Gracco, V. L., & De Nil, L. F.** (2015). The trajectory of gray matter development in Broca's area is abnormal in people who stutter. *Frontiers in Human Neuroscience*, 9, 89. <https://doi.org/10.3389/fnhum.2015.00089>
- Behzadi, Y., Restom, K., Liu, J., & Liu, T. T.** (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, 37(1), 90–101. <https://doi.org/10.1016/j.neuroimage.2007.04.042>
- Belyk, M., Kraft, S. J., & Brown, S.** (2015). Stuttering as a trait or state—An ALE meta-analysis of neuroimaging studies. *European Journal of Neuroscience*, 41(2), 275–284. <https://doi.org/10.1111/ejn.12765>
- Belyk, M., Kraft, S. J., & Brown, S.** (2017). Stuttering as a trait or a state revisited: Motor system involvement in persistent developmental stuttering. *European Journal of Neuroscience*, 45(4), 622–624. <https://doi.org/10.1111/ejn.13512>
- Bengtsson, S. L., Ehrsson, H. H., Forssberg, H., & Ullén, F.** (2004). Dissociating brain regions controlling the temporal and ordinal structure of learned movement sequences. *European Journal of Neuroscience*, 19(9), 2591–2602. <https://doi.org/10.1111/j.0953-816X.2004.03269.x>
- Bengtsson, S. L., Ehrsson, H. H., Forssberg, H., & Ullén, F.** (2005). Effector-independent voluntary timing: Behavioural and neuroimaging evidence. *European Journal of Neuroscience*, 22(12), 3255–3265. <https://doi.org/10.1111/j.1460-9568.2005.04517.x>
- 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(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Bernard, J. A., Peltier, S. J., Wiggins, J. L., Jaeggi, S. M., Buschkuhl, M., Fling, B. W., Kwak, Y., Jonides, J., Monk, C. S., & Seidler, R. D.** (2013). Disrupted cortico-cerebellar connectivity in older adults. *NeuroImage*, 83, 103–119. <https://doi.org/10.1016/j.neuroimage.2013.06.042>
- 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. <https://doi.org/10.1162/jocn.2009.21306>
- Bohland, J. W., & Guenther, F. H.** (2006). An fMRI investigation of syllable sequence production. *NeuroImage*, 32(2), 821–841. <https://doi.org/10.1016/j.neuroimage.2006.04.173>
- Braun, A. R., Varga, M., Stager, S., Schulz, G., Selbie, S., Maisog, J. M., Carson, R. E., & Ludlow, C. L.** (1997). Altered patterns of cerebral activity during speech and language production in developmental stuttering. An H2(15)O positron emission tomography study. *Brain*, 120(5), 761–784. <https://doi.org/10.1093/brain/120.5.761>
- 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. <https://doi.org/10.1002/hbm.20140>
- Buckner, R. L., Krienen, F. M., Castellanos, A., Diaz, J. C., & Yeo, B. T. T.** (2011). The organization of the human cerebellum estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106(5), 2322–2345. <https://doi.org/10.1152/jn.00339.2011>
- Bullmore, E. T., Suckling, J., Overmeyer, S., Rabe-Hesketh, S., Taylor, E., & Brammer, M. J.** (1999). Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain. *IEEE Transactions on Medical Imaging*, 18(1), 32–42. <https://doi.org/10.1109/42.750253>
- Cai, S., Tourville, J. A., Beal, D. S., Perkell, 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. <https://doi.org/10.3389/fnhum.2014.00054>
- Chang, S.-E., Erickson, K. I., Ambrose, N. G., Hasegawa-Johnson, M. A., & Ludlow, C. L.** (2008). Brain anatomy differences in childhood stuttering. *NeuroImage*, 39(3), 1333–1344. <https://doi.org/10.1016/j.neuroimage.2007.09.067>
- Chang, S.-E., & Guenther, F. H.** (2020). Involvement of the cortico-basal ganglia-thalamocortical loop in developmental stuttering. *Frontiers in Psychology*, 10, 3088. <https://doi.org/10.3389/fpsyg.2019.03088>
- Chang, S.-E., Horwitz, B., Ostuni, J., Reynolds, R., & Ludlow, C. L.** (2011). Evidence of left inferior frontal-premotor structural and functional connectivity deficits in adults who stutter. *Cerebral Cortex*, 21(11), 2507–2518. <https://doi.org/10.1093/cercor/bhr028>
- Chang, S.-E., Kenney, M. K., Loucks, T. M. J., & Ludlow, C. L.** (2009). Brain activation abnormalities during speech and non-speech in stuttering speakers. *NeuroImage*, 46(1), 201–212. <https://doi.org/10.1016/j.neuroimage.2009.01.066>
- Chang, S.-E., & Zhu, D. C.** (2013). Neural network connectivity differences in children who stutter. *Brain*, 136(12), 3709–3726. <https://doi.org/10.1093/brain/awt275>
- Chumbley, J., Worsley, K., Flandin, G., & Friston, K.** (2010). Topological FDR for neuroimaging. *NeuroImage*, 49(4), 3057–3064. <https://doi.org/10.1016/j.neuroimage.2009.10.090>
- Civier, O., Bullock, D., Max, L., & Guenther, F. H.** (2013). Computational modeling of stuttering caused by impairments in a basal ganglia-thalamo-cortical circuit involved in syllable selection and initiation. *Brain and Language*, 126(3), 263–278. <https://doi.org/10.1016/j.bandl.2013.05.016>
- Corbetta, M., & Shulman, G. L.** (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, 3(3), 201–215. <https://doi.org/10.1038/nrn755>
- Coull, J. T.** (2004). fMRI studies of temporal attention: Allocating attention within, or towards, time. *Cognitive Brain Research*, 21(2), 216–226. <https://doi.org/10.1016/j.cogbrainres.2004.02.011>
- Craig, A., Blumgart, E., & Tran, Y.** (2009). The impact of stuttering on the quality of life in adults who stutter. *Journal of Fluency Disorders*, 34(2), 61–71. <https://doi.org/10.1016/j.jfludis.2009.05.002>
- Craig, A., & Tran, Y.** (2006). Fear of speaking: Chronic anxiety and stammering. *Advances in Psychiatric Treatment*, 12(1), 63–68. <https://doi.org/10.1192/apt.12.1.63>
- Craig, A., & Tran, Y.** (2014). Trait and social anxiety in adults with chronic stuttering: Conclusions following meta-analysis. *Journal of Fluency Disorders*, 40, 35–43. <https://doi.org/10.1016/j.jfludis.2014.01.001>
- Davidow, J. H.** (2014). Systematic studies of modified vocalization: The effect of speech rate on speech production measures during metronome-paced speech in persons who stutter. *International Journal of Language & Communication Disorders*, 49(1), 100–112. <https://doi.org/10.1111/1460-6984.12050>
- De Nil, L. F., Beal, D. S., Lafaille, S. J., Kroll, R. M., Crawley, A. P., & Gracco, V. L.** (2008). The effects of simulated stuttering and prolonged speech on the neural activation patterns of stuttering and nonstuttering adults. *Brain and Language*, 107(2), 114–123. <https://doi.org/10.1016/j.bandl.2008.07.003>
- De Nil, L. F., Kroll, R. M., & Houle, S.** (2001). Functional neuroimaging of cerebellar activation during single word reading

- and verb generation in stuttering and nonstuttering adults. *Neuroscience Letters*, 302(2–3), 77–80. [https://doi.org/10.1016/S0304-3940\(01\)01671-8](https://doi.org/10.1016/S0304-3940(01)01671-8)
- De Nil, L. F., Kroll, R. M., Kapur, S., & Houle, S. (2000). A positron emission tomography study of silent and oral single word reading in stuttering and nonstuttering adults. *Journal of Speech, Language, and Hearing Research*, 43(4), 1038–1053. <https://doi.org/10.1044/jslhr.4304.1038>
- De Nil, L. F., Kroll, R. M., Lafaille, S. J., & Houle, S. (2003). A positron emission tomography study of short- and long-term treatment effects on functional brain activation in adults who stutter. *Journal of Fluency Disorders*, 28(4), 357–380. <https://doi.org/10.1016/j.jfludis.2003.07.002>
- Diedrichsen, J. (2006). A spatially unbiased atlas template of the human cerebellum. *NeuroImage*, 33(1), 127–138. <https://doi.org/10.1016/j.neuroimage.2006.05.056>
- Diedrichsen, J., Balsters, J. H., Flavell, J., Cussans, E., & Ramnani, N. (2009). A probabilistic MR atlas of the human cerebellum. *NeuroImage*, 46(1), 39–46. <https://doi.org/10.1016/j.neuroimage.2009.01.045>
- Diedrichsen, J., Maderwald, S., Küper, M., Thürling, M., Rabe, K., Gizewski, E. R., Ladd, M. E., & Timmann, D. (2011). Imaging the deep cerebellar nuclei: A probabilistic atlas and normalization procedure. *NeuroImage*, 54(3), 1786–1794. <https://doi.org/10.1016/j.neuroimage.2010.10.035>
- Duffy, J. R. (2013). *Motor speech disorders: Substrates, differential diagnosis, and management* (3rd ed.). Elsevier.
- Eden, G. F., Joseph, J. E., Brown, H. E., Brown, C. P., & Zeffiro, T. A. (1999). Utilizing hemodynamic delay and dispersion to detect fMRI signal change without auditory interference: The behavior interleaved gradients technique. *Magnetic Resonance in Medicine*, 41(1), 13–20. [https://doi.org/10.1002/\(SICI\)1522-2594\(199901\)41:1<13::AID-MRM4>3.0.CO;2-T](https://doi.org/10.1002/(SICI)1522-2594(199901)41:1<13::AID-MRM4>3.0.CO;2-T)
- Etchell, A. C., Civier, O., Ballard, K. J., & Sowman, P. F. (2018). A systematic literature review of neuroimaging research on developmental stuttering between 1995 and 2016. *Journal of Fluency Disorders*, 55, 6–45. <https://doi.org/10.1016/j.jfludis.2017.03.007>
- Etchell, A. C., Johnson, B. W., & Sowman, P. F. (2014). Behavioral and multimodal neuroimaging evidence for a deficit in brain timing networks in stuttering: A hypothesis and theory. *Frontiers in Human Neuroscience*, 8, 467. <https://doi.org/10.3389/fnhum.2014.00467>
- Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis. *NeuroImage*, 9(2), 195–207. <https://doi.org/10.1006/nimg.1998.0396>
- Fox, P. T., Ingham, R. J., Ingham, J. C., Hirsch, T. B., Downs, J. H., Martin, C., Jerabek, P., Glass, T., & Lancaster, J. L. (1996). A PET study of the neural systems of stuttering. *Nature*, 382(6587), 158–162. <https://doi.org/10.1038/382158a0>
- Fox, P. T., Ingham, R. J., Ingham, J. C., Zamarripa, F., Xiong, J.-H., & Lancaster, J. L. (2000). Brain correlates of stuttering and syllable production: A PET performance-correlation analysis. *Brain*, 123(10), 1985–2004. <https://doi.org/10.1093/brain/123.10.1985>
- Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *NeuroImage*, 6(3), 218–229. <https://doi.org/10.1006/nimg.1997.0291>
- Garnett, E. O., Chow, H. M., Nieto-Castañón, A., Tourville, J. A., Guenther, F. H., & Chang, S.-E. (2018). Anomalous morphology in left hemisphere motor and premotor cortex of children who stutter. *Brain*, 141(9), 2670–2684. <https://doi.org/10.1093/brain/awy199>
- Giraud, A. (2008). Severity of dysfluency correlates with basal ganglia activity in persistent developmental stuttering. *Brain and Language*, 104(2), 190–199. <https://doi.org/10.1016/j.bandl.2007.04.005>
- Golfinoopoulos, E., Tourville, J. A., Bohland, J. W., Ghosh, S. S., Nieto-Castanon, A., & Guenther, F. H. (2011). fMRI investigation of unexpected somatosensory feedback perturbation during speech. *NeuroImage*, 55(3), 1324–1338. <https://doi.org/10.1016/j.neuroimage.2010.12.065>
- Gracco, V. L., Tremblay, P., & Pike, B. (2005). Imaging speech production using fMRI. *NeuroImage*, 26(1), 294–301. <https://doi.org/10.1016/j.neuroimage.2005.01.033>
- Guenther, F. H. (2016). *Neural control of speech*. MIT Press. <https://doi.org/10.7551/mitpress/10471.001.0001>
- Guitar, B. (2014). *Stuttering: An integrated approach to its nature and treatment* (4th ed.). Lippincott Williams & Wilkins.
- 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. <https://doi.org/10.1016/j.neuroimage.2006.07.036>
- Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, A. Q., Elliott, M. R., Gurney, E. M., & Bowtell, R. W. (1999). “Sparse” temporal sampling in auditory fMRI. *Human Brain Mapping*, 7(3), 213–223. [https://doi.org/10.1002/\(SICI\)1097-0193\(1999\)7:3<213::AID-HBM5>3.0.CO;2-N](https://doi.org/10.1002/(SICI)1097-0193(1999)7:3<213::AID-HBM5>3.0.CO;2-N)
- Handwerker, D. A., Ollinger, J. M., & D’Esposito, M. (2004). Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *NeuroImage*, 21(4), 1639–1651. <https://doi.org/10.1016/j.neuroimage.2003.11.029>
- Hanna, R., & Morris, S. (1977). Stuttering, speech rate, and the metronome effect. *Perceptual and Motor Skills*, 44(2), 452–454. <https://doi.org/10.2466/pms.1977.44.2.452>
- Hashimoto, Y., & Sakai, K. L. (2003). Brain activations during conscious self-monitoring of speech production with delayed auditory feedback: An fMRI study. *Human Brain Mapping*, 20(1), 22–28. <https://doi.org/10.1002/hbm.10119>
- Haslinger, B., Erhard, P., Weilke, F., Ceballos-Baumann, A. O., Bartenstein, P., Gräfin von Einsiedel, H., Schwaiger, M., Conrad, B., & Boecker, H. (2002). The role of lateral premotor-cerebellar-parietal circuits in motor sequence control: A parametric fMRI study. *Cognitive Brain Research*, 13(2), 159–168. [https://doi.org/10.1016/S0926-6410\(01\)00104-5](https://doi.org/10.1016/S0926-6410(01)00104-5)
- Heim, S., Amunts, K., Hensel, T., Grande, M., Huber, W., Binkofski, F., & Eickhoff, S. B. (2012). The role of human parietal area 7A as a link between sequencing in hand actions and in overt speech production. *Frontiers in Psychology*, 3, 534. <https://doi.org/10.3389/fpsyg.2012.00534>
- Hickok, G., Buchsbaum, B., Humphries, C., & Muftuler, T. (2003). Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. *Journal of Cognitive Neuroscience*, 15(5), 673–682.
- Hutchinson, J. M., & Norris, G. M. (1977). The differential effect of three auditory stimuli on the frequency of stuttering behaviors. *Journal of Fluency Disorders*, 2(4), 283–293. [https://doi.org/10.1016/0094-730X\(77\)90032-8](https://doi.org/10.1016/0094-730X(77)90032-8)
- Ingham, R. J., Fox, P. T., Costello Ingham, J., & Zamarripa, F. (2000). Is overt stuttered speech a prerequisite for the neural activations associated with chronic developmental stuttering? *Brain and Language*, 75(2), 163–194. <https://doi.org/10.1006/brln.2000.2351>
- Institute of Electrical and Electronics Engineers. (1969). IEEE recommended practice for speech quality measurements. *IEEE Transactions on Audio and Electroacoustics*, 17, 227–246. <https://doi.org/10.1109/IEEESTD.1969.7405210>



- Ingham, R. J., Grafton, S. T., Bothe, A. K., & Ingham, J. C. (2012). Brain activity in adults who stutter: Similarities across speaking tasks and correlations with stuttering frequency and speaking rate. *Brain and Language*, 122(1), 11–24. <https://doi.org/10.1016/j.bandl.2012.04.002>
- Janssen, N., & Mendieta, C. C. R. (2020). The dynamics of speech motor control revealed with time-resolved fMRI. *Cerebral Cortex*, 30(1), 241–255. <https://doi.org/10.1093/cercor/bhz084>
- Jenkins, I., Brooks, D., Nixon, P., Frackowiak, R., & Passingham, R. (1994). Motor sequence learning: A study with positron emission tomography. *The Journal of Neuroscience*, 14(6), 3775–3790. <https://doi.org/10.1523/JNEUROSCI.14-06-03775.1994>
- Kell, C. A., Neumann, K., Behrens, M., von Gudenberg, A. W., & Giraud, A.-L. (2018). Speaking-related changes in cortical functional connectivity associated with assisted and spontaneous recovery from developmental stuttering. *Journal of Fluency Disorders*, 55, 135–144. <https://doi.org/10.1016/j.jfludis.2017.02.001>
- Kell, C. A., Neumann, K., von Kriegstein, K., Posenenske, C., von Gudenberg, A. W., Euler, H., & Giraud, A.-L. (2009). How the brain repairs stuttering. *Brain*, 132(10), 2747–2760. <https://doi.org/10.1093/brain/awp185>
- Keuken, M. C., Bazin, P.-L., Crown, L., Hootsmans, J., Laufer, A., Müller-Axt, C., Sier, R., van der Putten, E. J., Schäfer, A., Turner, R., & Forstmann, B. U. (2014). Quantifying inter-individual anatomical variability in the subcortex using 7 T structural MRI. *NeuroImage*, 94, 40–46. <https://doi.org/10.1016/j.neuroimage.2014.03.032>
- Krauth, A., Blanc, R., Poveda, A., Jeanmonod, D., Morel, A., & Székely, G. (2010). A mean three-dimensional atlas of the human thalamus: Generation from multiple histological data. *NeuroImage*, 49(3), 2053–2062. <https://doi.org/10.1016/j.neuroimage.2009.10.042>
- Lee, A., & Kawahara, T. (2009, January). *Recent development of open-source speech recognition engine Julius* [Paper presentation]. Asia-Pacific Signal and Information Processing Association 2009 Annual Summit and Conference, Sapporo, Japan.
- Lu, C., Chen, C., Ning, N., Ding, G., Guo, T., Peng, D., Yang, Y., Li, K., & Lin, C. (2010). The neural substrates for atypical planning and execution of word production in stuttering. *Experimental Neurology*, 221(1), 146–156. <https://doi.org/10.1016/j.expneurol.2009.10.016>
- Lu, C., Chen, C., Peng, D., You, W., Zhang, X., Ding, G., Deng, X., Yan, Q., & Howell, P. (2012). Neural anomaly and reorganization in speakers who stutter: A short-term intervention study. *Neurology*, 79(7), 625–632. <https://doi.org/10.1212/WNL.0b013e31826356d2>
- Lu, C., Ning, N., Peng, D., Ding, G., Li, K., Yang, Y., & Lin, C. (2009). The role of large-scale neural interactions for developmental stuttering. *Neuroscience*, 161(4), 1008–1026. <https://doi.org/10.1016/j.neuroscience.2009.04.020>
- Lu, C., Peng, D., Chen, C., Ning, N., Ding, G., Li, K., Yang, Y., & Lin, C. (2010). Altered effective connectivity and anomalous anatomy in the basal ganglia-thalamocortical circuit of stuttering speakers. *Cortex*, 46(1), 49–67. <https://doi.org/10.1016/j.cortex.2009.02.017>
- Max, L. (2004). Stuttering and internal models for sensorimotor control: A theoretical perspective to generate testable hypotheses. In B. Maassen, R. Kent, P. Hermann, & P. Van Lieshout (Eds.), *Speech motor control: In normal and disordered speech* (pp. 357–387). Oxford University Press.
- Max, L., Guenther, F. H., Gracco, V. L., Ghosh, S. S., & Wallace, M. E. (2004). Unstable or insufficiently activated internal models and feedback-biased motor control as sources of dysfluency: A theoretical model of stuttering. *Contemporary Issues in Communication Science and Disorders*, 31(Spring), 105–122. [https://doi.org/10.1044/cicsd\\_31\\_S\\_105](https://doi.org/10.1044/cicsd_31_S_105)
- McKay, L. C., Evans, K. C., Frackowiak, R. S. J., & Corfield, D. R. (2003). Neural correlates of voluntary breathing in humans. *Journal of Applied Physiology*, 95(3), 1170–1178. <https://doi.org/10.1152/jappphysiol.00641.2002>
- McLaren, D. G., Ries, M. L., Xu, G., & Johnson, S. C. (2012). A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. *NeuroImage*, 61(4), 1277–1286. <https://doi.org/10.1016/j.neuroimage.2012.03.068>
- Neumann, K., Euler, H. A., von Gudenberg, A. W., Giraud, A.-L., Lanfermann, H., Gall, V., & Preibisch, C. (2003). The nature and treatment of stuttering as revealed by fMRI: A within- and between-group comparison. *Journal of Fluency Disorders*, 28(4), 381–410. <https://doi.org/10.1016/j.jfludis.2003.07.003>
- Neumann, K., Preibisch, C., Euler, H. A., von Gudenberg, A. W., Lanfermann, H., Gall, V., & Giraud, A.-L. (2005). Cortical plasticity associated with stuttering therapy. *Journal of Fluency Disorders*, 30(1), 23–39. <https://doi.org/10.1016/j.jfludis.2004.12.002>
- Nieto-Castañón, A. (2020). *Handbook of functional connectivity Magnetic Resonance Imaging methods in CONN*. Hilbert Press.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1), 97–113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4)
- Parkinson, A. L., Flagmeier, S. G., Manes, J. L., Larson, C. R., Rogers, B., & Robin, D. A. (2012). Understanding the neural mechanisms involved in sensory control of voice production. *NeuroImage*, 61(1), 314–322. <https://doi.org/10.1016/j.neuroimage.2012.02.068>
- Pellegrino, F., Coupé, C., & Marsico, E. (2011). Across-language perspective on speech information rate. *Language*, 87(3), 539–558. <https://doi.org/10.1353/lan.2011.0057>
- Pellegrino, F., Farinas, J., & Rouas, J.-L. (2004). Automatic estimation of speaking rate in multilingual spontaneous speech. In *Proceedings of the Second International Conference on Speech Prosody* (pp. 517–520).
- Perrachione, T. K., & Ghosh, S. S. (2013). Optimized design and analysis of sparse-sampling fMRI experiments. *Frontiers in Neuroscience*, 7, 55. <https://doi.org/10.3389/fnins.2013.00055>
- Poldrack, R. A., Nichols, T., & Mumford, J. (2011). *Handbook of functional MRI data analysis*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511895029>
- Preibisch, C., Neumann, K., Raab, P., Euler, H. A., von Gudenberg, A. W., Lanfermann, H., & Giraud, A.-L. (2003). Evidence for compensation for stuttering by the right frontal operculum. *NeuroImage*, 20(2), 1356–1364. [https://doi.org/10.1016/S1053-8119\(03\)00376-8](https://doi.org/10.1016/S1053-8119(03)00376-8)
- Riecker, A., Kassubek, J., Gröschel, K., Grodd, W., & Ackermann, H. (Eds.). (2006). The cerebral control of speech tempo: Opposite relationship between speaking rate and BOLD signal changes at striatal and cerebellar structures. *NeuroImage*, 29(1), 46–53. <https://doi.org/10.1016/j.neuroimage.2005.03.046>
- Riley, G. D. (2008). *SSI-4, Stuttering Severity Instrument for Children and Adults—Fourth Edition*. Pro-Ed.
- Rottschy, C., Langner, R., Dogan, I., Reetz, K., Laird, A. R., Schulz, J. B., Fox, P. T., & Eickhoff, S. B. (2012). Modelling neural correlates of working memory: A coordinate-based meta-analysis. *NeuroImage*, 60(1), 830–846. <https://doi.org/10.1016/j.neuroimage.2011.11.050>
- Salmelin, R., Schnitzler, A., Schmitz, F., & Freund, H.-J. (2000). Single word reading in developmental stutterers and fluent speakers. *Brain*, 123(6), 1184–1202. <https://doi.org/10.1093/brain/123.6.1184>



- Schubotz, R. I., & von Cramon, D. Y.** (2001). Interval and ordinal properties of sequences are associated with distinct premotor areas. *Cerebral Cortex*, *11*(3), 210–222. <https://doi.org/10.1093/cercor/11.3.210>
- Segawa, J. A., Tourville, J. A., Beal, D. S., & Guenther, F. H.** (2015). The neural correlates of speech motor sequence learning. *Journal of Cognitive Neuroscience*, *27*(4), 819–831. [https://doi.org/10.1162/jocn\\_a\\_00737](https://doi.org/10.1162/jocn_a_00737)
- Sitek, K. R., Cai, S., Beal, D. S., Perkell, J. S., Guenther, F. H., & Ghosh, S. S.** (2016). Decreased cerebellar-orbitofrontal connectivity correlates with stuttering severity: Whole-brain functional and structural connectivity associations with persistent developmental stuttering. *Frontiers in Human Neuroscience*, *10*, 190. <https://doi.org/10.3389/fnhum.2016.00190>
- Stager, S. V., Denman, D. W., & Ludlow, C. L.** (1997). Modifications in aerodynamic variables by persons who stutter under fluency-evoking conditions. *Journal of Speech, Language, and Hearing Research*, *40*(4), 832–847. <https://doi.org/10.1044/jslhr.4004.832>
- Stager, S. V., Jeffries, K. J., & Braun, A. R.** (2003). Common features of fluency-evoking conditions studied in stuttering subjects and controls: An H<sub>2</sub><sup>15</sup>O PET study. *Journal of Fluency Disorders*, *28*(4), 319–336. <https://doi.org/10.1016/j.jfludis.2003.08.004>
- Takaso, H., Eisner, F., Wise, R. J. S., & Scott, S. K.** (2010). The effect of delayed auditory feedback on activity in the temporal lobe while speaking: A positron emission tomography study. *Journal of Speech, Language, and Hearing Research*, *53*(2), 226–236. [https://doi.org/10.1044/1092-4388\(2009/09-0009\)](https://doi.org/10.1044/1092-4388(2009/09-0009))
- Tourville, J. A., Reilly, K. J., & Guenther, F. H.** (2008). Neural mechanisms underlying auditory feedback control of speech. *NeuroImage*, *39*(3), 1429–1443. <https://doi.org/10.1016/j.neuroimage.2007.09.054>
- Toyomura, A., Fujii, T., & Kuriki, S.** (2011). Effect of external auditory pacing on the neural activity of stuttering speakers. *NeuroImage*, *57*(4), 1507–1516. <https://doi.org/10.1016/j.neuroimage.2011.05.039>
- Toyomura, A., Fujii, T., & Kuriki, S.** (2015). Effect of an 8-week practice of externally triggered speech on basal ganglia activity of stuttering and fluent speakers. *NeuroImage*, *109*, 458–468. <https://doi.org/10.1016/j.neuroimage.2015.01.024>
- Van Borsel, J., Achten, E., Santens, P., Lahorte, P., & Voet, T.** (2003). fMRI of developmental stuttering: A pilot study. *Brain and Language*, *85*(3), 369–376. [https://doi.org/10.1016/S0093-934X\(02\)00588-6](https://doi.org/10.1016/S0093-934X(02)00588-6)
- Vossel, S., Geng, J. J., & Fink, G. R.** (2014). Dorsal and ventral attention systems: Distinct neural circuits but collaborative roles. *The Neuroscientist*, *20*(2), 150–159. <https://doi.org/10.1177/1073858413494269>
- Watkins, K. E., Smith, S. M., Davis, S., & Howell, P.** (2007). Structural and functional abnormalities of the motor system in developmental stuttering. *Brain*, *131*(1), 50–59. <https://doi.org/10.1093/brain/awm241>
- Whitfield-Gabrieli, S., & Nieto-Castanon, A.** (2012). Conn: A functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connectivity*, *2*(3), 125–141. <https://doi.org/10.1089/brain.2012.0073>
- Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., & Evans, A. C.** (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping*, *4*(1), 58–73. [https://doi.org/10.1002/\(SICI\)1097-0193\(1996\)4:1<58::AID-HBM4>3.0.CO;2-O](https://doi.org/10.1002/(SICI)1097-0193(1996)4:1<58::AID-HBM4>3.0.CO;2-O)
- Yairi, E., & Ambrose, N. G.** (1999). Early childhood stuttering I: Persistency and recovery rates. *Journal of Speech, Language, and Hearing Research*, *42*(5), 1097–1112. <https://doi.org/10.1044/jslhr.4205.1097>
- Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fischl, B., Liu, H., & Buckner, R. L.** (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *106*(3), 1125–1165. <https://doi.org/10.1152/jn.00338.2011>

---

## Appendix A

### Stimulus Sentences Used in the Present Experiment

---

1. Rice is often served in round bowls.
  2. The juice of lemons makes fine punch.
  3. The boy was there when the sun rose.
  4. Her purse was full of useless trash.
  5. Hoist the load to your left shoulder.
  6. The young girl gave no clear response.
  7. Sickness kept him home the third week.
  8. Lift the square stone over the fence.
  9. The friendly gang left the drug store.
  10. The lease ran out in sixteen weeks.
  11. The steady bat gave birth to pups.
  12. There are more than two factors here.
  13. The lawyer tried to lose his case.
  14. The term ended late June that year.
  15. The pipe began to rust while new.
  16. Act on these orders with great speed.
- 

## Appendix B

### Speaking Instructions for Participants

---

During an earlier behavioral study outside the scanner, subjects were shown a PowerPoint presentation that included the following instructions:

“In this experiment, we will ask you to read aloud short sentences in two different ways:

- a rhythmic way, paced by a regular beat in the earphones you will wear
- a normal (non-rhythmic) way.”

“At the beginning of each trial, before you start reading, you will hear eight beats.

Those beats will always be regular.”

“In trials of nonrhythmic (normal) speech, the font will be (green/blue), and there will be the word ‘Normal’ above the sentence. Speak normally in these trials.”

“In trials of rhythmic speech, the font will be (blue/green), and there will be the word ‘Rhythm’ above the sentence. Speak rhythmically by aligning each syllable (vowel) to a beat.”

Prior to the scanning session, they were told the following:

“In the second part of the study, you will read sentences either in a rhythmic way or in a natural way. The crosshair (+) is your cue to stop reading. If you feel that you have said the sentence incorrectly, please do not ‘go back’ and try to correct it. Always keep your head and body as still as possible even while reading the sentences. On some trials, instead of sentences, you will see characters you cannot read. During these trials, please look at the characters and keep your head and body as still as possible.”

---