doi:10.1093/brain/awy199 BRAIN 2018: 141; 2670–2684 | 2670

BRAIN
A JOURNAL OF NEUROLOGY

Anomalous morphology in left hemisphere motor and premotor cortex of children who stutter

Emily O. Garnett,[1] Ho Ming Chow,[2] Alfonso Nieto-Castañón,[3] Jason A. Tourville,[3] Frank H. Guenther[3,4] and Soo-Eun Chang[1]

Stuttering is a neurodevelopmental disorder that affects the smooth flow of speech production. Stuttering onset occurs during a dynamic period of development when children first start learning to formulate sentences. Although most children grow out of stuttering naturally, ~1% of all children develop persistent stuttering that can lead to significant psychosocial consequences throughout one's life. To date, few studies have examined neural bases of stuttering in children who stutter, and even fewer have examined the basis for natural recovery versus persistence of stuttering. Here we report the first study to conduct surface-based analysis of the brain morphometric measures in children who stutter. We used FreeSurfer to extract cortical size and shape measures from structural MRI scans collected from the initial year of a longitudinal study involving 70 children (36 stuttering, 34 controls) in the 3–10-year range. The stuttering group was further divided into two groups: persistent and recovered, based on their later longitudinal visits that allowed determination of their eventual clinical outcome. A region of interest analysis that focused on the left hemisphere speech network and a whole-brain exploratory analysis were conducted to examine group differences and group × age interaction effects. We found that the persistent group could be differentiated from the control and recovered groups by reduced cortical thickness in left motor and lateral premotor cortical regions. The recovered group showed an age-related decrease in local gyrification in the left medial premotor cortex (supplementary motor area and and pre-supplementary motor area). These results provide strong evidence of a primary deficit in the left hemisphere speech network, specifically involving lateral premotor cortex and primary motor cortex, in persistent developmental stuttering. Results further point to a possible compensatory mechanism involving left medial premotor cortex in those who recover from childhood stuttering.

1 Department of Psychiatry, University of Michigan, Ann Arbor, MI, USA
2 Nemours/Alfred I. DuPont Hospital for Children, Wilmington, DE, USA
3 Department of Speech Language and Hearing Sciences, Boston University, Boston, MA, USA
4 Department of Biomedical Engineering, Boston University, Boston, MA, USA

Correspondence to: Soo-Eun Chang, PhD
Department of Psychiatry, University of Michigan
Rachel Upjohn Building, 4250 Plymouth Rd.
Ann Arbor, MI 48109, USA
E-mail: sooeunc@med.umich.edu

Keywords: cortical thickness; FreeSurfer; local gyrification index (LGI); MRI; surface area

Abbreviations: (GO)DIVA = (gradient order) directions into velocities of articulators; IFG = inferior frontal gyrus; LGI = local gyrification index; PMC = premotor cortex; SMA = supplementary motor area; SSI-4 = Stuttering Severity Instrument-4

Received March 16, 2018. Revised May 10, 2018. Accepted June 4, 2018. Advance Access publication July 31, 2018

Introduction

Developmental stuttering is a childhood onset neurodevelopmental disorder that affects 1% of the general population. At its core, stuttering is a speech disorder characterized by frequently occurring involuntary disruptions such as sound/syllable and word repetitions, prolongations, and blocking of sounds that severely impede the fluent flow of speech production. Stuttering is linked to both structural and functional abnormalities in brain regions involved in motor control and timing of speech movements. One convergent finding from previous investigations points to anomalous function and anatomy in left hemisphere structures involved in speech production (referred to here as the speech network). For instance, left motor cortical regions that mediate speech planning and production, including the inferior frontal gyrus (IFG) and the adjacent ventral premotor cortex (ventral PMC), were found to exhibit abnormal developmental trajectories in grey matter volume (Cykowski *et al.*, 2008; Kell *et al.*, 2009; Beal *et al.*, 2015), increased cortical folding (Foundas *et al.*, 2001), decreased underlying white matter integrity (Fox *et al.*, 2000; Sommer *et al.*, 2002; Chang *et al.*, 2009; Cykowski *et al.*, 2010), and reduced cerebral blood flow (Desai *et al.*, 2017; Neef *et al.*, 2018) in stuttering speakers. In addition, relative to controls, stuttering speakers exhibit decreased structural (Chang *et al.*, 2008, 2009; Watkins *et al.*, 2008; Cykowski *et al.*, 2010; Beal *et al.*, 2013; Chang and Zhu, 2013; Cai *et al.*, 2014*b*; Kronfeld-Duenias *et al.*, 2016) and functional connectivity (Lu *et al.*, 2010; Chang *et al.*, 2011; Qiao *et al.*, 2017) involving the left IFG/ventral PMC and other brain areas (e.g. the posterior superior temporal gyrus) of the speech network that support fluent speech production.

In addition to anomalous anatomy and function of left hemisphere cortical structures, some right hemisphere homologues have been found to exhibit greater structural volume, greater number of gyral banks/cortical folding [Cykowski *et al.*, 2008; however, Foundas *et al.*, (2001) found greater cortical folding in both left and right sylvian opercula], heightened functional activity (Fox *et al.*, 2000; Chang *et al.*, 2009), and greater structural connectivity (Neef *et al.*, 2018) in stuttering speakers.

The studies summarized above were all performed on adults who stutter. However, stuttering is a developmental disorder that starts in childhood (typically around 2–4 years of age), and it is well known that those whose stuttering persists develop secondary behaviours that complicate interpretation of findings involving adults who stutter. To date there have been only a handful of neuroimaging studies of children who stutter, and like most of the adult studies mentioned above, many of these studies have relied on statistical tests that did not involve rigorous correction for the large number of voxel-based comparisons involved in a whole-brain analysis. Although the lack of statistical correction forces caution when considering the results of these studies because of the high potential for false positives, they provide an important foundation for generating hypotheses to 'narrow the search area' for subsequent studies, which in turn allows for more definitive conclusions based on properly corrected statistics. The most common finding across morphometric studies of children who stutter (Chang *et al.*, 2008; Beal *et al.*, 2013; Chang and Zhu, 2013) is 'anomalous structure within the left hemisphere speech network'. For example, an early study by Chang *et al.* (2008) used voxel-based morphometry to compare grey matter volume in children who stutter and fluent children. The largest differences in grey matter volume were found in left IFG and left precentral gyrus (which includes motor and PMC); these areas are both crucial centres in the speech production network (Guenther, 2006). Children who stutter had smaller grey matter volume in these areas than controls. Beal *et al.* (2013) also found smaller inferior frontal gyrus grey matter volume in children who stutter compared to fluent children, although in this study the differences were found bilaterally. Chang and Zhu (2013) found structural differences between children who stutter and fluent children in white matter tracts primarily within the left hemisphere speech network, including connections between putamen, auditory cortex, supplementary material area (SMA), and insula. Analyses of resting state functional connectivity in the same subjects largely corroborated the tractography results. Chow and Chang (2017), in the first longitudinal study of childhood stuttering, showed that children who stutter had significant decreases in white matter integrity along the left arcuate fasciculus (a major white matter tract that interconnects the motor and auditory regions of the left hemisphere speech network). Children who stutter also exhibited decreased white matter integrity in corpus callosum areas containing fibres that interconnect the bilateral motor and auditory cortices (Chow and Chang, 2017). Further, this study found that children who continue to stutter versus those who recover from stuttering could be differentiated by distinct developmental trajectories; compared to the recovered group, who showed normalized growth with age, the persistent children showed stagnant white matter integrity increases with age in the left arcuate fasciculus, anterior thalamic radiation, and cerebral peduncles. Related, although not a study of morphology, Walsh *et al.* (2017) used functional near infrared spectroscopy (fNIRS) to examine cortical activity focused on bilateral IFG and superior temporal gyrus cortical areas during continuous speech production. The left IFG/ventral premotor region was the only region showing significantly aberrant patterns of the haemodynamic response during the speech production task in children who stutter compared to controls. The group differences in the right hemisphere homologues (IFG, superior temporal gyrus) were not significant (Walsh *et al.*, 2017).

Based on the prior studies of brain morphology in children who stutter cited above, we predicted that compared to fluent children, children with persistent stuttering would display morphological anomalies in the network of left

hemisphere cortical regions underlying speech production, and furthermore that children who recover from stuttering would show differences in morphology compared to children whose stuttering persists. Since the exact anatomical locations of anomalies noted in prior studies of stuttering have varied within the speech network, we included core sensorimotor regions as well as higher-order cortical areas involved in speech production (Guenther, 2016). Within this context, the purpose of the current study was to identify differences in brain morphology between fluent children (the control group), children who stuttered initially but recovered (the recovery group), and children whose stuttering persists (the persistent group). We focus here on morphology of the cerebral cortex, including measures of the size and local gyrification of cortical regions of interest.

The current study extends beyond prior work in several ways. First, we analysed data from a large paediatric sample spanning preschool to school-age children (3–10 year olds at initial testing). This allowed us, for the first time, to examine cortical morphology differences encompassing children close to stuttering onset. Second, we compare cortical morphometry of children who start out stuttering but eventually recover to children whose stuttering persists. This is possible because the data are part of a longitudinal study that tracks fluency and brain morphometry of children who stutter over the course of several years. Third, we characterize changes in morphometry as a function of age in childhood stuttering and in typically developing children. Fourth, we use image processing and statistical analysis methods that provide increased sensitivity to group differences in morphology than those used in prior studies, including cortical surface reconstruction (Dale *et al.*, 1999) and a functional-anatomical parcellation of cerebral cortex designed specifically for studies of speech that accounts for individual differences in cortical anatomy, thereby providing increased statistical power (Nieto-Castanon *et al.*, 2003; Tourville and Guenther, 2012). Fifth, this is the first study to investigate local gyrification in children. As noted above, this morphometric feature has been shown to differ in speech-related cortical areas of adults who stutter compared to fluent adults.

Materials and methods

Participants

Participants included 70 children (36 children who stutter, 14 females; 34 controls, 17 females) between 37.1 and 129.2 months of age. Demographic information for the two groups can be found in Table 1. All participants were right-handed monolingual speakers of English. Children were scanned up to four times (one visit per year) as part of a longitudinal study of brain morphometry and function in children who stutter; here we report cross-sectional data using scans from each child's first session. Scans were obtained from 87 participants, with 17 of those participants being removed from the participant pool for the current study because of image quality issues (primarily motion-based) that prevented the extraction of cortical surfaces (16 participants) or morphometric measures (one participant) using the FreeSurfer analysis software.

Participants completed a battery of standardized speech, language, and cognitive tests. They received audiometric hearing screening, oral-motor screening, and cognitive evaluations, details of which can be found in Chang *et al.* (2015). Children with scores below two standard deviations (SD) of the mean on any of the standardized assessments were excluded.

Stuttering severity was assessed using samples of spontaneous speech tasks with a parent and a certified speech-language pathologist. We calculated the percentage of disfluent syllables based on narrative samples and a monologue (storytelling) using a pictures-only book ['Frog, where are you?' (Mayer, 1969)]. In addition, the Stuttering Severity Instrument (SSI-4; Riley, 2009) was used to examine the frequency and duration of disfluencies occurring in the speech sample, as well as any physical concomitants associated with moments of stuttering; all of these measures were incorporated into a composite stuttering severity rating. To determine measurement reliability of the SSI score ratings, an intraclass correlation coefficient (ICC) was calculated based on two independent judges' ratings of SSI from a random subset (~44%) of the children's speech samples. The ICC for the overall SSI measurement between two independent judges was 0.98.

All children were trained during a separate visit with a mock MRI scanner to familiarize them with the MRI environment and procedures, and to practice keeping still while lying down inside the bore for stretches of time. Recordings of MRI scanning noises were played during this session, so that children were aware that they would be hearing loud MRI sounds during scanning. This session was repeated in some children, as needed. All procedures used in this study were approved by the Michigan State University Institutional Review Board. Informed consent was obtained according to the Declaration of Helsinki. All children were paid a nominal remuneration, and were given small prizes (e.g. stickers) for their participation.

While all children who stuttered were diagnosed with stuttering during their initial visit, they were later categorized as recovered or persistent through a combination of measures acquired in subsequent visits. Specifically, a child was considered recovered if the composite SSI-4 score was <10 at the second visit or thereafter. A child was categorized as persistent if the SSI-4 score was $\geqslant 10$ (corresponding to 'very mild' in SSI-4 severity classification) at the second visit or thereafter, and the onset of stuttering had been at least 36 months prior to the most recent visit. Determination of recovery status also required the consideration of per cent occurrence of stuttering-like disfluencies (%SLD) in the speech sample (>3 for persistent) as well as clinician and parental reports. Similar criteria were used to determine persistence versus recovery in stuttering children in previous studies (Yairi *et al.*, 1996; Yairi and Ambrose, 1999). Using these criteria, we identified 11 children who recovered and 25 children with persistent stuttering in the final dataset for the analyses. For controls, the inclusion criteria included: never diagnosed as stuttering, no family history of stuttering, lack of parental concern for their child's fluency, and a %SLD <3. A total of 34 controls were included.

Table 1 Demographics and behavioural scores for all participant groups

	Controls, *n* = 34 (17 boys)		Persistent, *n* = 25 (17 boys)		Recovered, *n* = 11 (5 boys)	
	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range
Age at initial visit	6.6 (2.0)	3.3–10.8	6.4 (1.9)	3.1–9.6	5.8 (2.1)	3.8–9.4
SES[a]	6.2 (0.5)	5.0–7.0	6.1 (0.8)	4.0–7.0	6.3 (0.6)	5.0–7.0
IQ[a,b]	115.7 (12.6)	87–144	106.1 (16.4)	78–138	107.9 (12.5)	88–128
PPVT[b]	117.7 (12.9)	95–151	107.1 (9.9)	90–131	114.91 (17.3)	85–147
EVT[b]	115.5 (13.9)	90–149	105.5 (12.3)	85–134	109.7 (10.2)	94–127
GFTA[c]	104.2 (8.3)	76–123	102.2 (6.9)	87–118	108.1 (7.2)	99–121
SSI-4 at initial visit[d]	N/A	N/A	22.7 (7.4)	12–48	16.1 (4.1)	11–22
SSI-4 at final visit[d]	N/A	N/A	20.3 (9.2)	7–48	8.3 (2.1)	4–11
Months post-onset[a]	N/A	N/A	38.2 (25.2)	6–90	24.1 (20.5)	7–70

[a]Tests measured only at each participant's initial visit.
[b]Scores significantly lower in persistent than controls (two-sample *t*-test, $P < 0.05$).
[c]Scores significantly lower in persistent than recovered (two-sample *t*-test, $P < 0.05$).
[d]Scores significantly higher in persistent than recovered (two-sample *t*-test, $P < 0.05$).
EVT = Expressive Vocabulary Test; GFTA-2 = Goldman-Fristoe Test of Articulation; IQ = intelligence quotient; N/A = not applicable; PPVT = Peabody Picture Vocabulary Test; SES = socioeconomic status.

MRI acquisition

All MRI scans were acquired on a GE 3 T Signa HDx MR scanner with an 8-channel head coil. During each session, whole brain T_1-weighted inversion recovery fast spoiled gradient-recalled images (3D IRFSPGR) with CSF suppressed were obtained with the following parameters: echo time = 3.8 ms, repetition time of acquisition = 8.6 ms, inversion time = 831 ms, repetition time of inversion = 2332 ms, flip angle = 8, field of view = 25.6 cm × 25.6 cm, matrix size = 256 × 256, slice thickness = 1 mm, and receiver bandwidth = ±20.8 kHz. The T_1-weighted images were acquired as part of a longitudinal imaging study that also included acquisition of DTI and resting state functional MRI data. Children viewed a movie, and a research staff member sat next to the child to ensure comfort and compliance throughout the scanning procedure (~40 min).

Image processing

FreeSurfer 5.3.0 (http://surfer.nmr.mgh.harvard.edu/) was used to automatically segment individual T_1-weighted anatomical volumes and to generate 3D reconstructions of each individual's cortical surface. The procedure included motion correction, intensity bias correction, skull stripping, and tissue classification. Triangular tessellation was then applied to create representations of white matter and pial surfaces. Image segmentation and surface reconstructions were visually inspected; when surface errors were present that were the result of poor image segmentation, manual edits were made in accordance with the FreeSurfer tutorial (https://surfer.nmr.mgh.harvard.edu/fswiki/FsTutorial/TroubleshootingData) and surface reconstructions were regenerated.

Following reconstruction, each subject's cortical surface was divided into 62 distinct anatomical regions (parcels) per hemisphere based on individual anatomical landmarks according to the SLaparc parcellation system (Tourville and Guenther, 2012). A representative cortical surface template was constructed from the surface reconstructions of 28 participants, including 14 children who stutter (seven female; mean age 82 months) and 14 controls (seven female; 88 months). The SLaparc parcellation system was mapped from the FreeSurfer adult *fsaverage* template to the representative paediatric surface template and an expert rater (J.T.) manually inspected and corrected the resulting labelled surface to ensure accurate adherence to the SLaparc parcellation system. Each individual surface reconstruction was co-registered to the representative template and the SLaparc labels were mapped from the template to the individual surface. Template generation, surface co-registration, and surface-to-surface label mapping were all completed with tools in the FreeSurfer 5.3.0 software distribution.

Five morphometric measures were extracted from FreeSurfer for each anatomical parcel: average cortical thickness, surface area, volume, thickness-to-area ratio, and local gyrification index (LGI).

Two types of analyses were run to detect group morphometry differences: a hypothesis-based analysis focused on finding expected morphometry differences between groups in the left hemisphere speech network with statistical corrections for the number of regions and morphometric measures (see 'Analyses of group differences' section below), and an exploratory analysis involving all 62 parcels per hemisphere with no statistical correction for the number of regions. The hypothesis-based analysis was limited to 26 of 62 left hemisphere parcels that have been identified as part of the speech production network based on prior functional neuroimaging studies (for review see Guenther, 2016). These regions were grouped into 14 functional regions of interest, each containing one to three anatomically-defined parcels from the SLaparc parcellation. Table 2 lists the set of functional regions of interest and corresponding anatomical parcels, and Fig. 1 illustrates the functional regions of interest and parcels on an inflated cortical surface. This 'nested' approach, which is an example of a hierarchical fixed-sequence testing procedure for multiple hypothesis testing (Bretz *et al.*, 2009), was used to maximize statistical power in subsequent analyses of group differences by utilizing subject-specific regions of interest based on expected function-anatomy associations (Nieto-Castanon *et al.*, 2003) while providing more precise localization of group differences within larger functional regions of interest that contain multiple anatomical subregions. The functional regions of interest used here include

Table 2 Functional regions of interest and corresponding anatomical parcels

Functional regions of interest (n = 14)	Anatomical parcels (n = 26)
Lateral premotor cortex	vPMC, midPMC
Medial premotor cortex	SMA, preSMA
Primary motor cortex	midMC, aCO, vMC
Primary somatosensory cortex	pCO, vSC
Higher order somatosensory cortex	PO, aSMg
Primary auditory cortex	Hg
Anterior higher order auditory cortex	PP, aSTg, adSTs
Posterior higher order auditory cortex	pSTg, pdSTs, PT
Inferior frontal gyrus pars opercularis	IFo, pFO
Inferior frontal gyrus pars triangularis	IFt, aFO
Inferior frontal sulcus	pIFs
Frontal orbital cortex	FOC
Posterior supramarginal gyrus	pSMg
Anterior insula	aINS

See Tourville and Guenther (2012) for details regarding anatomical landmarks delineating anatomical parcels. aCO = anterior central operculum; adSTs = anterior dorsal superior temporal sulcus; aINS = anterior insula; aFO = anterior frontal operculum; aSMg = anterior supramarginal gyrus; aSTg = anterior superior temporal gyrus convexity; FOC = frontal orbital cortex; Hg = Heschl's gyrus; IFo = inferior frontal gyrus pars opercularis convexity; IFt = inferior frontal gyrus pars triangularis convexity; midMC = middle motor cortex; midPMC = middle premotor cortex; pCO = posterior central operculum; pdSTs = posterior dorsal superior temporal sulcus; pFO = posterior frontal operculum; pIFs = posterior inferior frontal sulcus; PO = parietal operculum; PP = planum polare; pSMg = posterior supramarginal gyrus; pSTg = posterior superior temporal gyrus convexity; PT = planum temporale; vMC = ventral motor cortex; vPMC = ventral PMC; vSC = ventral somatosensory cortex.

core sensorimotor areas (primary motor cortex, medial PMC, lateral PMC, primary somatosensory cortex, higher order somatosensory cortex, primary auditory cortex, anterior higher order auditory cortex, posterior higher order auditory cortex) as well as association and paralimbic regions that have been implicated in speech production (inferior frontal gyrus pars opercularis, inferior frontal gyrus pars triangularis, inferior frontal sulcus, frontal orbital cortex, posterior supramarginal gyrus, and anterior insula); see Guenther (2016) for hypothesized functions of these regions in speech.

Morphometric measure selection

We investigated two aspects of cortical morphology in separate analyses: region of interest size and gyrification. For the gyrification analysis, the FreeSurfer measure LGI, which characterizes the amount of cortex within sulcal folds compared to the outer cortex, was the dependent variable. For the region of interest size analysis, we first performed a dimensionality reduction analysis using all four FreeSurfer measures of region of interest size (cortical thickness, surface area, volume, and thickness-to-area ratio). This analysis was motivated by redundancies in the region of interest size measures, which significantly reduce statistical power if all measures are included in an analysis involving statistical correction for multiple comparisons. To ameliorate this potential problem, the four FreeSurfer size measures were submitted to a principal component analysis after being converted to z-scores separately for each of the 62 anatomical parcels per hemisphere, and then concatenated across all subjects and parcels. We found that the

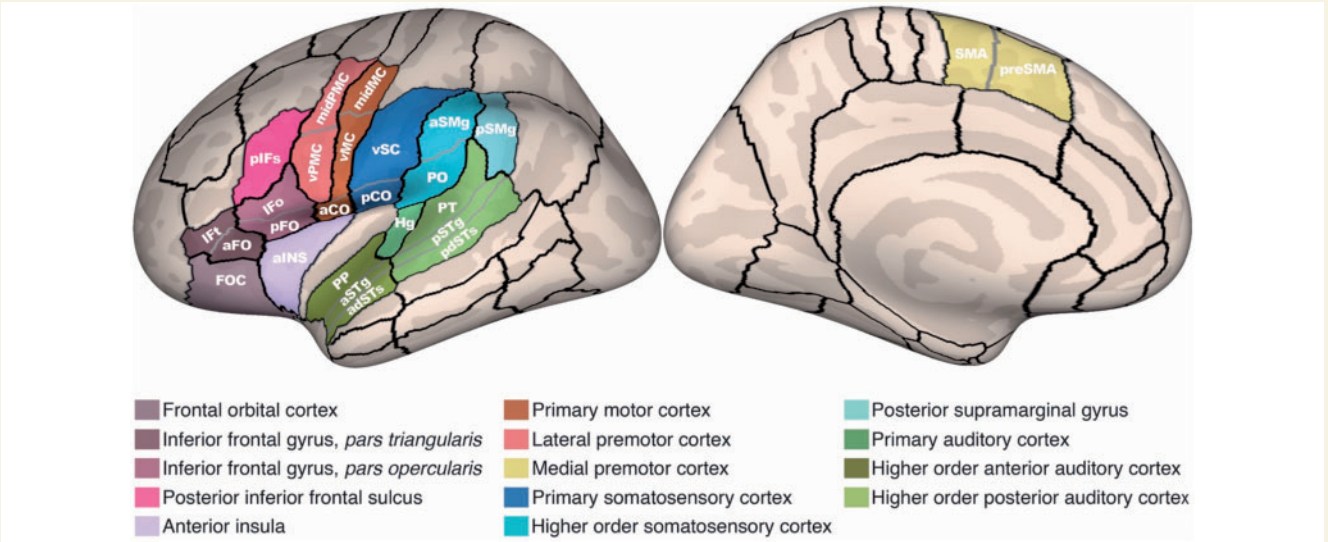

Figure 1 Functional regions of interest. The functional regions of interest and associated anatomical parcels are shown on an inflated reconstruction of a representative left hemisphere cortical surface. aCO = anterior central operculum; adSTs = anterior dorsal superior temporal sulcus; aINS = anterior insula; aFO = anterior frontal operculum; aSMg = anterior supramarginal gyrus; aSTg = anterior superior temporal gyrus convexity; FOC = frontal orbital cortex; Hg = Heschl's gyrus; IFo = inferior frontal gyrus pars opercularis convexity; IFt = inferior frontal gyrus pars triangularis convexity; midMC = middle motor cortex; midPMC = middle premotor cortex; pCO = posterior central operculum; pdSTs = posterior dorsal superior temporal sulcus; pFO = posterior frontal operculum; pIFs = posterior inferior frontal sulcus; PO = parietal operculum; PP = planum polare; pSMg = posterior supramarginal gyrus; pSTg = posterior superior temporal gyrus convexity; PT = planum temporale; vMC = ventral motor cortex; vSC = ventral somatosensory cortex.

first two principal components explained 98% of the variance. We then determined how much of the variance of the original four measures could be captured by each possible combination of two measures. Among all possible pairs, surface area and cortical thickness together explained the most variance (96%) across the original four variables, compared to 98% for the first two principle components. Based on these analyses, we chose to use surface area and cortical thickness as our two dependent variables for subsequent region of interest size analysis since (i) they account for the vast majority of the variance in the original four measures in a non-redundant fashion; and (ii) they are more straightforward to interpret than the first two principal components.

Covariate analysis of demographic factors

The potential confounding influence of demographic factors was addressed by performing a multivariate regression to identify significant covariation between four demographic variables (age, sex, IQ, and socioeconomic status) and three morphometric outcome measures (cortical thickness, surface area, LGI) and three morphometric outcome measures (cortical thickness, surface area, LGI) aggregated across all anatomical parcels. This analysis was limited to the set of 34 control participants. In addition, ANOVA was used to identify potential differences in demographic factors between the three subject groups. Demographic variables that showed statistically significant covariation with the outcome measures or significant group differences were used as control covariates in subsequent analyses. Because language test scores are highly correlated with IQ, scores for the Peabody Picture Vocabulary Test and Expressive Vocabulary Test were not included in these analyses to avoid multicollinearity issues. The Goldman-Fristoe Test of Articulation was also excluded as it may reflect greater articulatory variability that is associated with stuttering (St. Louis and Hinzman, 1988; Louko *et al.*, 1990; Wolk *et al.*, 1993; Blood *et al.*, 2003; Melnick *et al.*, 2003), particularly in persistent relative to recovered children who stutter (Paden *et al.*, 1999; Usler *et al.*, 2017).

Analyses of group differences

The chosen morphometric measures for region of interest size and gyrification were submitted to analyses of covariance (ANCOVA). Main effects of group and age × group interactions were examined. Age, sex, and IQ were also included as control covariates in the analyses. The outcome measures for the cortical size analysis were surface area and cortical thickness. LGI was the outcome measure for the cortical gyrification analysis. The contrasts of interest focusing on group differences included main group effects and group × age interactions. In other words, the analyses used F-tests to evaluate the presence of differences between groups in the cross-sectional developmental profiles of the outcome measures of interest, irrespective of whether these differences were linked to differences in the average levels of the outcome measure within each group (main group effects) or whether they were linked to differences in the strength of age-related changes in the outcome measure within each group (group × age interactions).

The hypothesis-based ANCOVA analysis involved the 14 left hemisphere speech network functional regions of interest (containing a total of 26 anatomical parcels). For cortical size analyses, we first performed an omnibus test across all 26 anatomical parcels in the 14 functional regions of interest to identify which individual measure(s) (surface area, cortical thickness) showed significant group differences (two separate ANCOVAs, one per measure), using false discovery rate (FDR) to correct for multiple comparisons across these two measures. For each significant measure found in this analysis, we then performed an ANCOVA analysis to identify significant group differences within each of the 14 functional regions of interest (14 separate ANCOVAs, one per functional region of interest), using FDR to correct for multiple comparisons across these 14 region of interests. Finally, for each significant measure and functional region of interest combination, we identified significant group differences within the anatomical parcels that comprise the functional region of interest (a variable number of ANCOVAs, one per anatomical parcel within each functional region of interest), using FDR to correct for multiple comparisons across these parcels. For the gyrification analysis, we performed an ANCOVA analysis to identify significant group differences in LGI within each of the 14 functional regions of interest (14 separate ANCOVAs, one per functional region of interest), using FDR to correct for multiple-comparisons across these functional regions of interest, then identified significant group differences within the anatomical parcels that comprise the functional region of interest (a variable number of ANCOVAs, one per anatomical parcel within each functional region of interest), using FDR to correct for multiple comparisons across these parcels.

The exploratory analysis involved all 62 anatomical parcels in each hemisphere and followed the above steps except that (i) no statistical corrections were applied for multiple comparisons across regions; and (ii) a single multivariate test (MANCOVA) involving all three morphometric measures (surface area, cortical thickness, LGI) was used to identify significant group differences within each functional region of interest. A second exploratory MANCOVA was performed to identify group differences in left-right asymmetry, in the form of a laterality index computed as (left − right) / (left + right), for each of the three morphometric measures.

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Results

Demographic factors

A multiple regression aimed at identifying covariation between demographic factors (age, sex, IQ, and Speaker's Experience of Stuttering) and morphometric outcome measures (cortical thickness, surface area, and LGI) in the control group found significant effects of age [$F(3,27) = 10.34$, $P = 0.0001$] and sex [$F(3,27) = 4.02$, $P = 0.017$]. These demographic measures were thus included as control covariates in subsequent between-group analyses. Older subjects exhibited greater surface area and lower cortical thickness when

compared to younger subjects. Females exhibited greater surface area and LGI compared to males. Group differences in demographic measures were identified by entering these variables as dependent measures in an ANOVA with group as independent factor. This revealed significant differences between groups in IQ [$F(2,68) = 3.66$, $P = 0.0309$], with controls exhibiting a higher mean IQ (115.7) than the persistent (106.0) and recovered (107.9) groups. IQ was thus included as an additional control covariate in subsequent analyses.

Hypothesis-based group analyses

Group effect analyses were performed on the 14 left hemisphere speech network functional regions of interest and corresponding anatomical parcels as described in the 'Materials and methods' section. Separate analyses were performed for region of interest size (with outcome measures cortical thickness and surface area) and gyrification (with outcome measure LGI).

Omnibus tests across individual region of interest size measures within all left hemisphere functional regions of interest lumped into a single region revealed significant group effects for cortical thickness [$F(4,62) = 3.02$, $P = 0.024$, $P_{FDR} = 0.048$] but not for surface area [$F(4,62) = 0.89$, $P = 0.474$], thereby supporting our primary hypothesis of group differences in morphometry within the left hemisphere speech network, in particular for cortical thickness. Subsequent analysis steps for region of interest size were thus performed only on cortical thickness in left hemisphere speech functional regions of interest. ANCOVA analysis of group differences within individual functional regions of interest revealed significant group effects for cortical thickness in two of the functional regions of interest in the left hemisphere: lateral PMC [$\chi^2(8) = 22.13$, $P = 0.005$, $P_{FDR} = 0.035$], and primary motor cortex [$\chi^2(12) = 28.27$, $P = 0.005$, $P_{FDR} = 0.035$]. *Post hoc* analysis of these functional regions of interest identified significant group cortical thickness differences in four anatomical parcels within these functional regions of interest: middle PMC [$F(4,62) = 3.55$, $P = 0.011$, $P_{FDR} = 0.026$], ventral motor cortex [$F(4,62) = 3.20$, $P = 0.019$, $P_{FDR} = 0.026$], ventral PMC [$F(4,62) = 3.17$, $P = 0.020$, $P_{FDR} = 0.026$], and for anterior central operculum [$F(4,62) = 3.13$, $P = 0.021$, $P_{FDR} = 0.026$]. Finally, *post hoc* analyses discriminating between main group effects and age × group interactions for cortical thickness within these four anatomical parcels allowed us to characterize the effects within those regions as follows:

(i) Middle PMC group effects were dominated by main between-group differences [$F(2,62) = 6.57$, $P = 0.003$, $P_{FDR} = 0.005$], with the persistent group having lower cortical thickness compared to the recovery and control groups [$t(62) = 3.26$, $P = 0.002$, $P_{FDR} = 0.007$; Fig. 2A].

(ii) Ventral motor cortex group effects (Fig. 2B) were dominated by main between-group differences [$F(2,62) = 5.64$, $P = 0.006$, $P_{FDR} = 0.011$], with lower cortical thickness in the persistent compared to the recovery group [$t(62) = 3.36$, $P = 0.001$, $P_{FDR} = 0.005$].

(iii) Ventral PMC group effects (Fig. 2C) were dominated by group × age interactions [$F(2,62) = 4.29$, $P = 0.018$, $P_{FDR} = 0.036$], with cortical thickness decreasing with age in the persistent group but not the control group [$t(62) = -2.881$, $P = 0.005$, $P_{FDR} = 0.022$],

(iv) Anterior central operculum group effects (Fig. 2D) were driven by a combination of both main between-group differences [$F(2,62) = 3.16$, $P = 0.049$, $P_{FDR} = 0.098$] and group × age interactions [$F(2,62) = 2.37$, $P = 0.102$, $P_{FDR} = 0.102$], but no significant individual effects.

The functional region of interest analysis for gyrification revealed significant group effects in one left hemisphere functional region of interest only: medial PMC [$\chi^2(8) = 24.36$, $P = 0.002$, $P_{FDR} = 0.028$]. *Post hoc* analysis identified significant group LGI differences in both anatomical parcels within this functional region of interest: SMA [$F(4,62) = 3.80$, $P = 0.008$, $P_{FDR} = 0.016$]; and preSMA [$F(4,62) = 3.09$, $P = 0.022$, $P_{FDR} = 0.022$]. Finally, *post hoc* analyses discriminating main group effects and age × group interactions for LGI within these parcels identified both significant main and interaction effects in SMA [main effect $F(2,62) = 6.027$, $P = 0.004$, $P_{FDR} = 0.008$; interaction effect $F(2,62) = 3.55$, $P = 0.035$, $P_{FDR} = 0.035$] consistent with a decrease in LGI with age in the recovery group but not in the persistent or control groups [$t(62) = -2.43$, $P = 0.018$, $P_{FDR} = 0.036$; Fig. 2E], as well as a significant main effect in preSMA [$F(2,62) = 5.794$, $P = 0.005$, $P_{FDR} = 0.010$] consistent with reduced LGI in the recovery group compared to the persistent or control groups [$t(62) = -3.25$, $P = 0.002$; Fig. 2F].

A summary of the significant group differences from the hypothesis-based analyses plotted on an inflated cortical surface is provided in Fig. 3.

Exploratory group analyses

Exploratory analyses used MANCOVA to identify potential main group effects or group × age interactions across the three subject groups (control, persistent, and recovery) within any of the three outcome measures (surface area, cortical thickness, and LGI). These analyses were performed separately within each parcel across a total of 124 parcels covering both hemispheres. Three parcels survived a threshold of $P < 0.01$ (uncorrected), suggesting potential effects pending replication: left SMA, right posterior parahippocampal gyrus, and left posterior ventral superior temporal sulcus. In left SMA, group effects [$\chi^2(12) = 30.43$, $P = 0.002$] were driven mainly by LGI differences [$F(4,62) = 3.80$, $P = 0.008$], consistent with the results observed in our main confirmatory analyses (namely, a decrease in LGI with age for the recovered group but not the other two groups; Fig. 2F). (Note that the lack of any significant left hemisphere

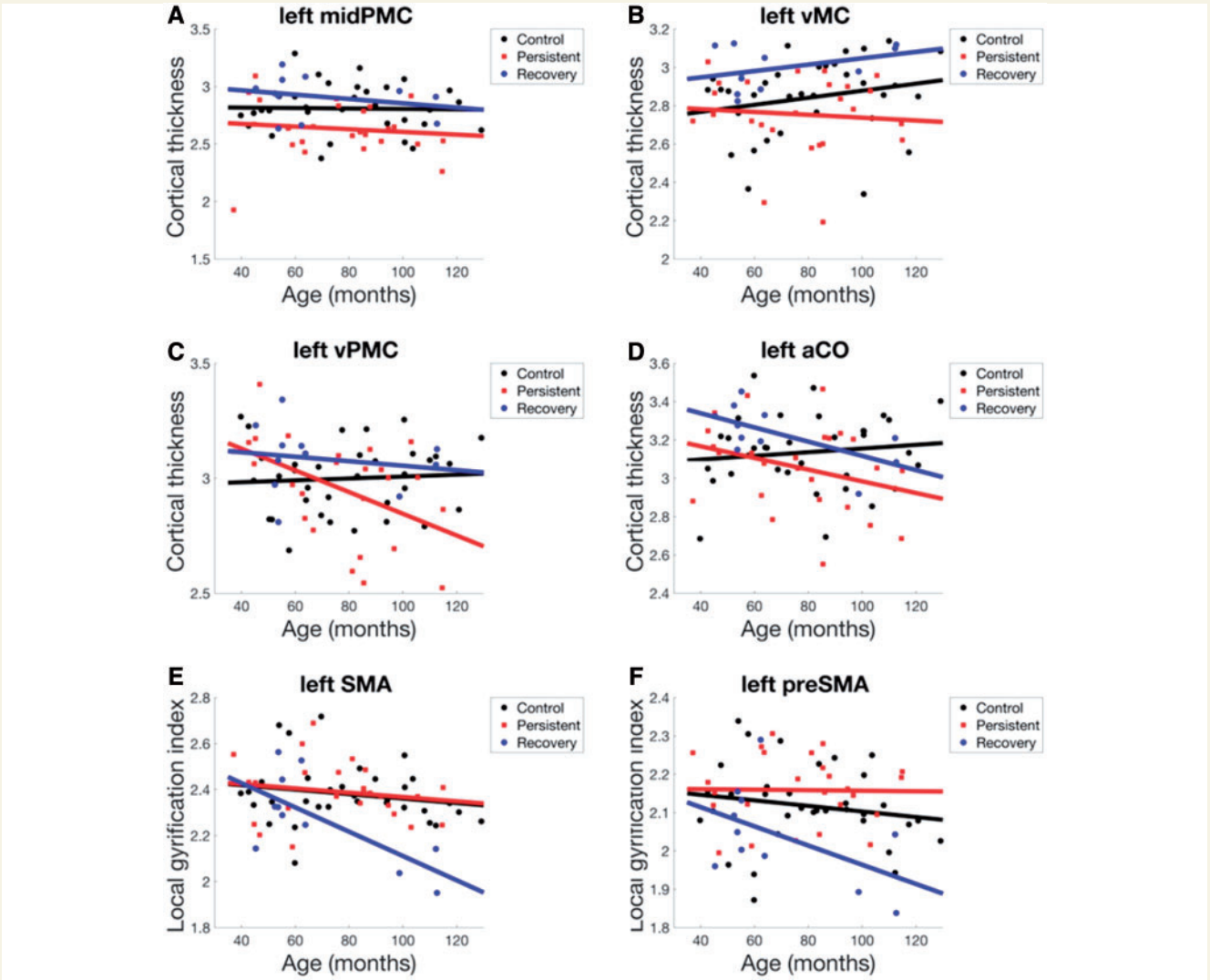

Figure 2 Premotor, motor, and medial motor cortical areas showing significant group differences in morphometry. Significant morphometric group differences (*P-FDR* < 0.05) were identified in ANCOVA analyses of group differences in left hemisphere speech network morphology, plotted as a function of age. aCO = anterior central operculum; midPMC = middle premotor cortex; preSMA = pre-supplementary motor area; vMC = ventral motor cortex; vPMC = ventral premotor cortex.

speech network differences beyond those identified in the nested hypothesis-based analyses indicates that the use of larger functional regions of interest in the hypothesis-based analyses did not mask group differences that may have been apparent in only one of the anatomical parcels within the functional region of interest.) Right posterior parahippocampal gyrus effects were also driven by differences in LGI [$F(4,62) = 4.11$, $P = 0.005$], with LGI trending upward with age in the control group but downward in the other two groups (Fig. 4A). Left left posterior ventral superior temporal sulcus effects appeared to be driven by differences in cortical thickness [$F(4,62) = 4.67$, $P = 0.002$], with cortical thickness trending more strongly downward with age in the persistent group than the other two groups (Fig. 4B).

It is possible that the uncorrected *P*-value threshold of $P < 0.01$ used for the exploratory analysis was less sensitive than the FDR-corrected threshold of 0.05 used in the hypothesis-based analysis stream, which in turn might explain why we found several left hemisphere anomalies but no right hemisphere anomalies in the speech network. To eliminate this possibility, we applied the hypothesis-based analysis stream to the right hemisphere speech network and still found no right hemisphere group differences.

A second exploratory analysis was performed to look for possible group differences in left/right asymmetry by calculating laterality indices for the three morphometric measures (surface area, cortical thickness, and LGI) and submitting them to a MANCOVA analysis as described above. Only one region, Heschl's gyrus, exhibited a significant group difference at the $P < 0.01$ uncorrected threshold [$\chi^2(12) = 30.01$, $P = 0.003$], driven primarily by differences in surface area asymmetry. As illustrated in Fig. 4C, the persistent and recovery groups had higher left-right surface area asymmetry in Heschl's gyrus compared to the control

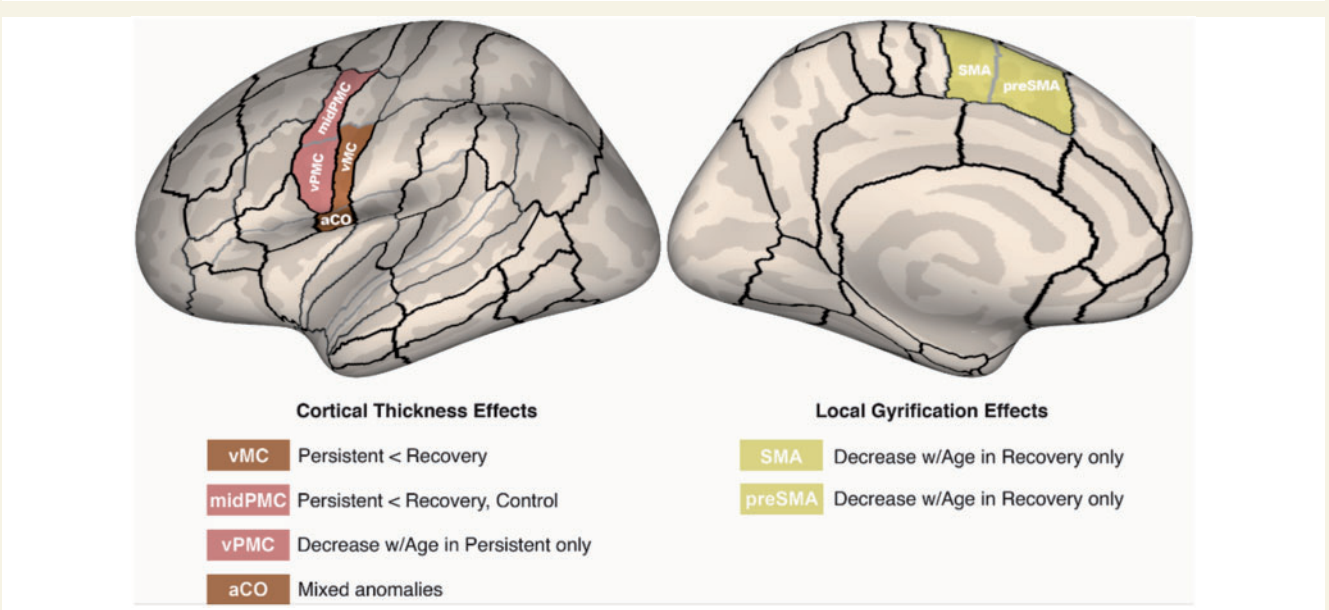

Figure 3 Summary of significant group differences in left hemisphere cortical morphology. Areas showing significant group differences are plotted on an inflated cortical surface template. aCO = anterior central operculum; midPMC = middle premotor cortex; preSMA = pre-supplementary motor area; vMC = ventral motor cortex; vPMC = ventral premotor cortex.

group. Furthermore, the recovery and control groups show a tendency for this asymmetry to increase with age, whereas asymmetry for the persistent group remains flat across age.

Discussion

In this study, we sought to determine if grey matter morphology in speech-related brain regions, as well as its developmental trajectory (determined cross-sectionally), would distinguish between three groups of young children: those who recovered from stuttering, those with persistent stuttering, and controls with no history of stuttering. To this end, we used a statistically-sensitive functional region of interest-based analysis of the cortical ribbon to derive and compare morphometric measures including cortical thickness, surface area, volume, thickness-to-area ratio, and LGI across groups. Preliminary analyses indicated that the vast majority of the variability in the data could be captured with three of these measures: the size measures cortical thickness and surface area, and the gyrification measure LGI. Subsequent analyses identified several group differences in cortical thickness and LGI in left hemisphere motor and premotor areas that are involved in speech production.

Based on previous findings reporting reduced grey matter volume in children who stutter in structures in the left hemisphere speech network, we expected to find focal cortical thickness decreases in these regions for children who stutter relative to controls. The current results showed this to be the case for children with persistent stuttering specifically: there was significantly decreased cortical thickness in left premotor and primary motor areas in the persistent stuttering group compared to the other groups. In the lateral PMC, the persistent stuttering group had lower cortical thickness in the ventral PMC compared to the control group and in the middle PMC compared to both recovery and control groups. The persistent group also showed decreased cortical thickness with age in ventral PMC that was not evident in the other groups. In primary motor cortex, we found lower cortical thickness in ventral motor cortex for persistent compared to recovered stuttering group. An effect of group was also found for anterior central operculum, which is adjacent to ventral motor cortex and forms part of the Rolandic operculum, an area where previous studies of adults (Sommer *et al.*, 2002), and children who stutter (Chang *et al.*, 2008) reported decreased integrity of underlying white matter tracts. Although no significant pairwise group differences were found in the anterior central operculum, the data were consistent with the ventral motor cortex finding of lower cortical thickness in persistent compared to recovered children (Fig. 2D).

The results further showed significant group differences with LGI in left medial premotor cortical areas. This included a decrease in LGI with age in SMA of the recovery group but not the control or persistent groups (Fig. 2E), and lower LGI in preSMA in the recovery group compared to the control and persistent groups (Fig. 2F). No significant group effects or interactions between group and age were found for surface area.

Although the neural deficits underlying stuttering are still an active topic of debate, many researchers have posited that the core deficit in stuttering is in the left hemisphere

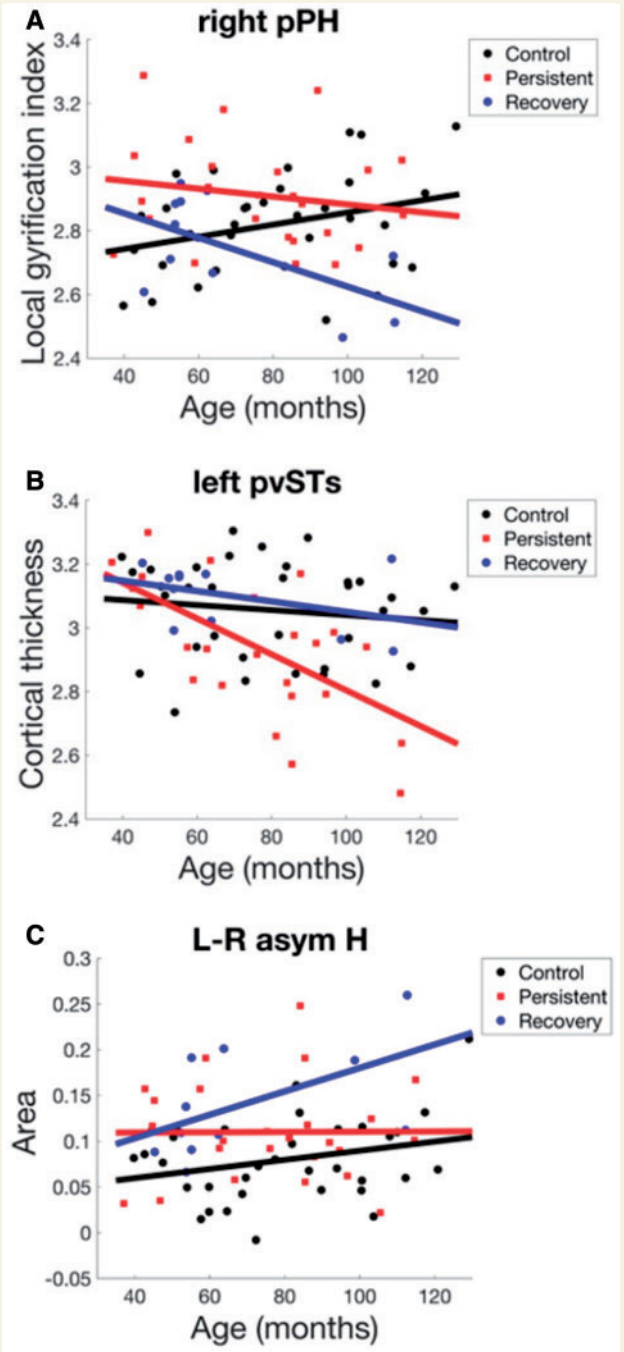

Figure 4 Areas with morphometric differences identified in the exploratory analyses ($P < 0.01$, uncorrected), plotted as a function of age. H = Heschl's gyrus; L-R asym = left/right asymmetry (laterality index); pPH = posterior parahippocampal gyrus; pvSTs = posterior ventral superior temporal sulcus.

basal ganglia-thalamo-cortical motor circuit (Maguire *et al.*, 2000, 2002, 2004; Alm, 2004; Chang and Zhu, 2013; Civier *et al.*, 2013), which we will refer to here as the basal ganglia motor loop. It should be noted that we did not directly assess potential structural differences in the basal ganglia region itself, given our focus on cortical measures, but rather discuss the cortical findings relevant to the basal ganglia motor loop network. The brain areas in which we found structural anomalies in children who stutter are all key components of this circuit (Middleton and Strick, 2000), which has been implicated in the selection and initiation of motor acts within behavioural sequences (Brotchie *et al.*, 1991; Marsden and Obeso, 1994; Mink, 1996) including the sequence of gestures for a word or syllable (Bohland and Guenther, 2006; Bohland *et al.*, 2010).

According to the DIVA neurocomputational model of speech motor control (directions into velocities of articulators) (Guenther, 2006; Guenther *et al.*, 2006), neurons in left hemisphere ventral PMC represent well-learned speech sequences such as frequently produced syllables in a 'speech sound map', and activation of a syllable's representation in this map leads to the readout of a finely tuned motor program for the syllable via projections to ventral motor cortex, cerebellum (via the pons), and the basal ganglia motor loop. The GODIVA model (gradient order directions into velocities of articulators) (Bohland *et al.*, 2010; Civier *et al.*, 2013) is an extension of DIVA that describes the neural circuitry underlying speech sound sequencing and motor program initiation. The DIVA/GODIVA framework accounts for a wide range of behavioural and neural findings concerning speech sequencing, and stuttering can be induced in computer simulations of the GODIVA model by impairing the basal ganglia motor loop (Civier *et al.*, 2013). In GODIVA, the basal ganglia motor loop is responsible for initiating the articulatory gestures within a syllabic motor program at the right instants in time by activating neurons in an 'initiation map' in SMA. Projections from sensory, motor, and premotor cortical areas to the putamen provide a detailed 'sensorimotor context' that the basal ganglia monitors to determine exactly when to initiate the next gesture in the sequence. For example, left ventral PMC provides information about the syllable currently being produced, SMA and ventral motor cortex provide information about the ongoing articulatory gesture, ventral somatosensory cortex provides information about the current somatosensory state, and posterior auditory cortex provides information about the current acoustic signal being produced. When the basal ganglia recognize that the current gesture is nearly complete, a 'completion signal' is sent to SMA that extinguishes activity in the initiation map neurons coding the current gesture and activates the neurons coding the next gesture.

Consideration of our morphometry results within the DIVA/GODIVA theoretical framework leads to the following interpretations. Lower cortical thickness in ventral motor cortex and ventral PMC in children who stutter may be indicative of impaired neural processing in these areas, which in turn makes it relatively difficult for the basal ganglia motor loop to identify the proper sensorimotor context for initiating the next gesture in a speech sequence, leading to moments of stuttering. Alternatively, reduced cortical thickness in ventral motor cortex and

ventral PMC may be a secondary consequence of impaired neural processing within the basal ganglia or SMA that leads to less effective activation of motor programs in ventral PMC/ventral motor cortex, and this reduced activity in turn leads to thinner cortex through some currently unknown neurodevelopmental process. The fact that significant differences found in persistent stuttering children were primarily in early developing cortical morphology (cortical thickness) provides some support for the former possibility. Future research that incorporates longitudinal modelling, as well as combined analysis of functional and structural MRI data may further elucidate this issue.

Our findings of group differences in SMA/preSMA LGI are more difficult to interpret. Cortical gyrification during the postnatal period shows peak growth between 2–6 years of age (Raznahan *et al.*, 2011) with generally protracted decreases in 6-year-olds and older. Cortical gyrification supports expansion of surface area, and it has been shown that increased LGI links to better cortical function such as higher intelligence (Luders *et al.*, 2005). On the other hand, higher mean levels of gyrification was found in children with autism relative to controls across the 4–12-year range, with abnormal age-related gyrification increases in the autism group (Yang *et al.*, 2016; see also Hardan *et al.*, 2004; Jou *et al.*, 2010). Greater gyrification is also linked to local short-range hyper-connectivity in children with autism (Schaer *et al.*, 2013). In addition, LGI was negatively correlated with more years of training (e.g. expert versus untrained divers; Guenther, 2016; Zhang *et al.*, 2016), indicating that decreasing LGI might represent synaptic pruning to support efficient neural circuitry supporting optimal function behaviours (White *et al.*, 2010; Li *et al.*, 2014).

Given that the persistent stuttering group did not show any LGI difference in SMA/preSMA compared to controls (anomalies were found only in the recovery group), it seems unlikely that this finding represents a root cause of stuttering. Instead, a decrease in LGI with age in the recovery group suggests that changes in left SMA/preSMA function and/or structure may somehow offset the neural processing impairments that caused these children to stutter when they were younger. Left SMA/preSMA is interconnected with the left posterior IFG via the frontal aslant tract (Dick *et al.*, 2014), which was shown in recent studies to support language production (Catani *et al.*, 2013). More specific to stuttering, Kronfeld-Duenias *et al.* (2016) reported that the left frontal aslant tract exhibited greater mean diffusivity in stuttering speakers relative to controls, and that left frontal aslant tract mean diffusivity values were negatively correlated with speech rate in stuttering speakers. The authors argued that increased mean diffusivity could have stemmed from '...a noisy communication (reduced synchrony) between IFG and SMA...' and that lower mean diffusivity values '...predict faster transmission between inferior frontal language regions and the preSMA/SMA involved in speech planning and production' (p. 378). In another study, Kemerdere and colleagues (2016) showed with axonal stimulation of frontal aslant tract, which provides a transient virtual lesion to the stimulated area, that disruption of left frontal aslant tract led to transitory stuttering (Kemerdere *et al.*, 2016). These studies suggest the critical role of the left frontal aslant tract in fluent speech production. If increased gyral folding is linked to increased connectivity of short tracts interconnecting local areas (Ecker *et al.*, 2016), lessening of the LGI in the SMA/preSMA may indicate synaptic pruning and fine-tuning of neural circuits involving this region. Namely, we speculate that decreased LGI with age in the left SMA/preSMA in the recovery group may underlie better long-range connectivity between left SMA/preSMA and left IFG that helps achieve more fluent speech. Future studies that combine examination of LGI and DTI tractography in stuttering children (persistent, recovered), would help confirm these ideas.

Although our exploratory finding of cortical thickness anomalies in left posterior ventral superior temporal sulcus (a higher order auditory cortical area) must be interpreted with caution because of the use of uncorrected statistics, it is possible that impaired auditory input to the putamen from left posterior ventral superior temporal sulcus may contribute to difficulties in recognizing the proper sensorimotor context for initiating upcoming gestures. An intriguing alternative possibility is motivated by the observation that the persistent group starts out with similar left posterior ventral superior temporal sulcus cortical thickness to the control and recovery groups at around 40 months of age but the persistent group shows a decline in cortical thickness with age not seen in the other groups (Fig. 4B). It is well established that a number of auditory feedback manipulations (e.g. masking noise, frequency-shifted feedback, or delayed auditory feedback; Adams and Ramig, 1980; Andrews *et al.*, 1980; Stuart *et al.*, 2008; Ingham *et al.*, 2009, 2012; Saltuklaroglu *et al.*, 2009; Foundas *et al.*, 2013) can induce fluency in people who stutter, at least temporarily. These manipulations may work because they reduce the likelihood that the basal ganglia will detect a mismatch in the sensorimotor context (in the form of a mismatch between expected and actual auditory feedback) for initiating the next gesture in the sequence. Over time, the brains of people with persistent stuttering may (subconsciously) learn to inhibit auditory processing of their own speech, thereby reducing (but not eliminating) the likelihood of a moment of stuttering (Guenther, 2016). Support for this idea comes from studies investigating sensorimotor adaptation to auditory perturbations, which indicate that adults who stutter show reduced adaptation compared to controls (Nudelman *et al.*, 1992; Cai *et al.*, 2012, 2014*a*; Loucks *et al.*, 2012) whereas children who stutter show the same amount of adaptation as fluent children (Daliri *et al.*, 2017).

Because prior morphometry studies involving children who stutter consistently found anomalies in the left hemisphere speech areas (with right hemisphere findings in children who stutter being less consistent, though not absent; Beal *et al.*, 2013) we focused our hypothesis-based analysis

on the left hemisphere speech network, allowing us to use statistical tests that were rigorously corrected for multiple comparisons. Nonetheless, it is noteworthy that our exploratory analysis did not find any significant right hemisphere speech network anomalies, even with a *P*-value threshold that was not corrected for multiple comparisons. The exploratory analysis did find a significant group difference in right posterior parahippocampal gyrus, but this difference is difficult to interpret given that posterior parahippocampal gyrus is not generally considered to be a speech area. The posterior parahippocampal gyrus is part of the limbic system and has been linked to contextual associations (Aminoff *et al.*, 2013), including contextual cues in speech such as sarcasm (Rankin *et al.*, 2009). While this was an unexpected finding, the significance of social context for stuttering severity is well established (Yaruss and Quesal, 2006; Craig *et al.*, 2014). Thus, hypothesis-driven future studies of stuttering focused on the limbic system including the right posterior parahippocampal gyrus might be warranted. According to the DIVA/GODIVA framework, feedforward motor programs for speech sequences are essentially stored in left ventral PMC (as discussed above), whereas right hemisphere ventral PMC is more heavily involved in sensory feedback-based adjustments of the motor commands. This is consistent with our finding of only left hemisphere anomalies in children who stutter since stuttering is an impairment of the readout of stored motor programs.

In contrast to our finding of only left hemisphere morphology differences in children who stutter, morphometry studies of adults who stutter consistently find right hemisphere anomalies in the speech network, mostly in the form of larger region of interest sizes/thicknesses and stronger white matter tracts (Jancke *et al.*, 2004; Neef *et al.*, 2018), which contrasts sharply with the smaller region of interest sizes/thicknesses and weaker tracts found in the left hemisphere of children who stutter (Chang *et al.*, 2008, 2015; Beal *et al.*, 2013; Chow and Chang, 2017). The natural interpretation of this pattern of results within the DIVA/GODIVA framework is that the core deficit in stuttering is an impairment of the left hemisphere feedforward control system (and thus left hemisphere anomalies are found in both adults and children who stutter), and this deficit forces over-reliance on right hemisphere feedback control mechanisms, eventually leading to right hemisphere morphological changes seen in adults who stutter.

One additional finding from our exploratory analyses was a group difference in left–right asymmetry in Heschl's gyrus, which is the location of the primary auditory cortex. A prior study involving adults found reduced asymmetry in the planum temporale, an auditory cortical region immediately caudal to Heschl's gyrus, of adults who stutter compared to age-matched controls (Foundas *et al.*, 2001). However, a more recent study that included younger participants failed to replicate this planum temporale asymmetry difference (Gough *et al.*, 2018), and the current study's results do not support reduced asymmetry in children who stutter for either Heschl's gyrus (where we found increased asymmetry in children who stutter) or planum temporale (where we found no significant group differences). One possible reason for these apparently conflicting findings may be that asymmetry in planum temporale and/or Heschl's gyrus changes with age in different ways for children who stutter compared to controls. Tentative support for this view is found in Fig. 4C, which indicates that laterality of Heschl's gyrus surface area increased with age in control participants and recovered stutterers, while asymmetry in persistent stutterers remained constant across age. A similar pattern was found for grey matter density in planum temporale by Gough *et al.* (2018). Extrapolating into adulthood, this could lead to a situation where adults who stutter have decreased laterality compared to those who do not. At present this interpretation should be considered speculative given the exploratory nature of our asymmetry analysis; however, our results provide a strong rationale for a future study of auditory cortical asymmetries in stuttering using longitudinal data and/or a large cohort covering a larger age range.

While we have thus far applied the DIVA/GODIVA theoretical framework to guide interpretation of our current findings, other theoretical accounts could provide alternative explanations. The significant group differences in cortical thickness found in the medial PMC, for example, might be explained in the context of this region being involved in generating movements that result from internal as opposed to external cues. This is interesting in light of hypotheses proposing that stuttering may result from an internal timing deficit related to impairment of basal ganglia thalamocortical connections that leads to the inability to generate or maintain internally-paced movements such as fluent speech production (Alm, 2004; Etchell *et al.*, 2014; Wieland *et al.*, 2015).

A so-called rhythm perception and timing network (Grahn and Rowe, 2009) includes putamen, SMA, and PMC regions, which continue to be reported in neuroimaging and neurophysiological studies of stuttering (De Nil *et al.*, 2003; Neumann *et al.*, 2005; Giraud *et al.*, 2008; Kell *et al.*, 2009; Lu *et al.*, 2009, 2010; Toyomura *et al.*, 2011; Chang and Zhu, 2013; Civier *et al.*, 2013; Chang *et al.*, 2016). The present study focuses examination of surface-based cortical morphometric measures and thus we cannot comment on subcortical regions that form critical components of this network. The cortical regions that are heavily interconnected with the putamen, however, including ventral motor cortex, ventral PMC, medial PMC, differentiated the persistent stuttering children from the other groups. Further, age-related decreases in the gyrification measure LGI in the SMA/preSMA were found in the recovered group. The SMA and putamen form the 'main core timing network' (Merchant *et al.*, 2013), and significantly decreased functional connectivity between these areas has been found in stuttering children relative to controls (Chang and Zhu, 2013). The present finding

of age-related LGI decreases in left SMA/preSMA in the recovery group thus leads to an intriguing question: could recovery be supported not only through a better long-range connectivity between SMA/preSMA and the left IFG, but also with the putamen, a major node of the rhythm/timing network? Current research underway that combines morphometric measures with DTI in a longitudinal design will help us to answer this question.

In sum, we report the first morphometric study of childhood stuttering focused on surface based cortical measures. The children who would eventually persist in stuttering showed early differentiation from the control and the eventually recovered groups in cortical thickness in left motor and lateral premotor areas. These results corroborate findings of aberrant articulatory coordination and movement indices in children who stutter, particularly in boys who are more likely to persist in stuttering symptoms (Walsh *et al.*, 2015). The children who would eventually recover showed decreased gyrification in left SMA/preSMA, which we tentatively interpret as a possible indicator of improved long-range connectivity with other cortical and subcortical areas that may help achieve fluent speech production. These results provide novel information that contributes to our expanding knowledge base on the neural bases of stuttering and the possible basis for chronicity versus natural recovery from stuttering.

Acknowledgements

The authors wish to thank all the children and parents who have participated in this study. We also thank Kristin Hicks for her assistance in participant recruitment, behavioural testing, and help with MRI data collection, Scarlett Doyle for her assistance in MRI data acquisition, Barbara Holland for assistance with data quality control, and Ashley Larva and for her assistance in speech data analyses.

Funding

This study was supported by Award Numbers R01DC011277 (SC) and R01DC007683 (FG) from the National Institute on Deafness and other Communication Disorders (NIDCD). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIDCD or the National Institutes of Health.

References

Adams MR, Ramig P. Vocal characteristics of normal speakers and stutterers during choral reading. J Speech Hear Res 1980; 23: 457–69.

Alm PA. Stuttering and the basal ganglia circuits: a critical review of possible relations. J Commun Disord 2004; 37: 325–70.

Aminoff EM, Kveraga K, Bar M. The role of the parahippocampal cortex in cognition. Trends Cogn Sci 2013; 17: 379–90.

Andrews G, Guitar B, Howie P. Meta-analysis of the effects of stuttering treatment. J Speech Hear Disord 1980; 45: 287–307.

Beal DS, Gracco VL, Brettschneider J, Kroll RM, De Nil LF. A voxel-based morphometry (VBM) analysis of regional grey and white matter volume abnormalities within the speech production network of children who stutter. Cortex 2013; 49: 2151–61.

Beal DS, Lerch JP, Cameron B, Henderson R, Gracco VL, De Nil LF. The trajectory of gray matter development in Broca's area is abnormal in people who stutter. Front Hum Neurosci 2015; 9: 89.

Blood GW, Ridenour Jr VJ, Qualls CD, Hammer CS. Co-occurring disorders in children who stutter. J Commun Disord 2003; 36: 427–48.

Bohland JW, Bullock D, Guenther FH. Neural representations and mechanisms for the performance of simple speech sequences. J Cogn Neurosci 2010; 22: 1504–29.

Bohland JWJ, Guenther FHF. An fMRI investigation of syllable sequence production. Neuroimage 2006; 32: 821–41.

Bretz F, Maurer W, Brannath W, Posch M. A graphical approach to sequentially rejective multiple test procedures. Stat Med 2009; 28: 586–604.

Brotchie JM, Mitchell IJ, Sambrook MA. Alleviation of parkinsonism by antagonism of excitatory amino acid transmission in the medial segment of the globus pallidus in rat and primate. Mov Disord 1991; 6: 133–8.

Cai S, Beal DS, Ghosh SS, Guenther FH, Perkell JS. Impaired timing adjustments in response to time-varying auditory perturbation during connected speech production in persons who stutter. Brain Lang 2014a; 129: 24–9.

Cai S, Beal DS, Ghosh SS, Tiede MK, Guenther FH, Perkell JS. Weak responses to auditory feedback perturbation during articulation in persons who stutter: evidence for abnormal auditory-motor transformation. PLoS One 2012; 7: e41830–0.

Cai S, Tourville JA, Beal DS, Perkell JS, Guenther FH, Ghosh SS. Diffusion imaging of cerebral white matter in persons who stutter: evidence for network-level anomalies. Front Hum Neurosci 2014b; 8; 54.

Catani M, Mesulam MM, Jakobsen E, Malik F, Martersteck A, Wieneke C, et al. A novel frontal pathway underlies verbal fluency in primary progressive aphasia. Brain 2013; 136: 2619–28.

Chang S-E, Chow HM, Wieland EA, McAuley JD. Relation between functional connectivity and rhythm discrimination in children who do and do not stutter. Neuroimage Clin 2016; 12: 442–50.

Chang S-E, Erickson KI, Ambrose NG, Hasegawa-Johnson MA, Ludlow CL. Brain anatomy differences in childhood stuttering. Neuroimage 2008; 39: 1333–44.

Chang S-E, Horwitz B, Ostuni J, Reynolds R, Ludlow CL. Evidence of left inferior frontal-premotor structural and functional connectivity deficits in adults who stutter. Cereb Cortex 2011; 21: 2507–18.

Chang S-E, Kenney MK, Loucks TMJ, Ludlow CL. Brain activation abnormalities during speech and non-speech in stuttering speakers. Neuroimage 2009; 46: 201–12.

Chang S-E, Zhu DC, Choo AL, Angstadt M. White matter neuroanatomical differences in young children who stutter. Brain 2015; 138: 694–711.

Chang S-E, Zhu DC. Neural network connectivity differences in children who stutter. Brain 2013; 136: 3709–26.

Chow HM, Chang S-E. White matter developmental trajectories associated with persistence and recovery of childhood stuttering. Hum Brain Mapp 2017; 13: 266.

Civier O, Bullock D, Max L, Guenther FH. Computational modeling of stuttering caused by impairments in a Basal Ganglia Thalamocortical circuit involved in syllable selection and initiation. Brain Lang 2013; 126: 263–78.

Craig A. Major controversies in fluency disorders: clarifying the relationship between anxiety and stuttering. J Fluency Disord 2014; 40: 1–3.

Cykowski MD, Fox PT, Ingham RJ, Ingham JC, Robin DA. A study of the reproducibility and etiology of diffusion anisotropy

differences in developmental stuttering: a potential role for impaired myelination. Neuroimage 2010; 52: 1495–504.
Cykowski MD, Kochunov PV, Ingham RJ, Ingham JC, Mangin JF, Rivière D, et al. Perisylvian sulcal morphology and cerebral asymmetry patterns in adults who stutter. Cereb Cortex 2008; 18: 571–83.
Dale A, Fischl B, Sereno M. Cortical surface-based analysis: I. Segmentation and surface reconstruction. Neuroimage 1999; 9: 179–94.
Daliri A, Wieland EA, Cai S, Guenther FH, Chang S-E. Auditory-motor adaptation is reduced in adults who stutter but not in children who stutter. Dev Sci 2017; 121: e12521.
De Nil L, Kroll R, Lafaille S, Houle S. A positron emission tomography study of short- and long-term treatment effects on functional brain activation in adults who stutter. J Fluency Disord 2003; 28: 357–80.
Desai J, Huo Y, Wang Z, Bansal R, Williams SCR, Lythgoe D, et al. Reduced perfusion in Broca's area in developmental stuttering. Hum Brain Mapp 2017; 38: 1865–74.
Dick AS, Bernal B, Tremblay P. The language connectome: new pathways, new concepts. Neuroscientist 2014; 20: 453–67.
Ecker C, Andrews D, Dell'Acqua F, Daly E, Murphy C, Catani M, et al. Relationship between cortical gyrification, white matter connectivity, and autism spectrum disorder. Cereb Cortex 2016; 26: 3297–309.
Etchell AC, Johnson BW, Sowman PF. Beta oscillations, timing, and stuttering. Front Hum Neurosci 2014; 8: 1036.
Foundas AL, Bollich AM, Corey DM, Hurley M, Heilman KM. Anomalous anatomy of speech-language areas in adults with persistent developmental stuttering. Neurology 2001; 57: 207–15.
Foundas AL, Mock JR, Corey DM, Golob EJ, Conture EG. Brain and Language. Brain Lang 2013; 126: 141–50.
Fox PT, Ingham RJ, Ingham JC, Zamarripa F, Xiong J-H, Lancaster JL. Brain correlates of stuttering and syllable production A PET performance-correlation analysis. Brain 2000; 123: 1985–2004.
Giraud A, Neumann K, Bachoud-Levi A, Gudenberg von A, Euler H, Lanfermann H, et al. Severity of dysfluency correlates with basal ganglia activity in persistent developmental stuttering. Brain Lang 2008; 104: 190–9.
Gough PM, Connally EL, Howell P, Ward D, Chesters J, Watkins KE. Planum temporale asymmetry in people who stutter. J Fluency Disord 2018; 55: 94–105.
Grahn JA, Rowe JB. Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. J Neurosci 2009; 29: 7540–8.
Guenther F. Cortical interactions underlying the production of speech sounds. J Commun Disord 2006; 39: 350–65.
Guenther FH, Ghosh SS, Tourville JA. Neural modeling and imaging of the cortical interactions underlying syllable production. Brain Lang 2006; 96: 280–301.
Guenther FH. Neural control of speech. Cambridge, MA: MIT Press; 2016.
Hardan AY, Jou RJ, Keshavan MS, Varma R, Minshew NJ. Increased frontal cortical folding in autism: a preliminary MRI study. Psychiatry Res 2004; 131: 263–8.
Ingham RJ, Bothe AK, Jang E, Yates L, Cotton J, Seybold I. Measurement of speech effort during fluency-inducing conditions in adults who do and do not stutter. J Speech Lang Hear Res 2009; 52: 1286–301.
Ingham RJ, Bothe AK, Wang Y, Purkhiser K, New A. Phonation interval modification and speech performance quality during fluency-inducing conditions by adults who stutter. J Commun Disord 2012; 45: 198–211.
Jancke L, Hänggi J, Steinmetz H. Morphological brain differences between adult stutterers and non-stutterers. BMC Neurol 2004; 4: 23.
Jou RJ, Minshew NJ, Keshavan MS, Hardan AY. Cortical gyrification in autistic and Asperger disorders: a preliminary magnetic resonance imaging study. J Child Neurol 2010; 25: 1462–7.
Kell CA, Neumann K, Kriegstein Von K, Posenenske C, Gudenberg Von AW, Euler H, et al. How the brain repairs stuttering. Brain 2009; 132: 2747–60.
Kemerdere R, de Champfleur NM, Deverdun J, Cochereau J, Moritz-Gasser S, Herbet G, et al. Role of the left frontal aslant tract in stuttering: a brain stimulation and tractographic study. J Neurol 2016; 263: 157–67.
Kronfeld-Duenias V, Amir O, Ezrati-Vinacour R, Civier O, Ben-Shachar M. Dorsal and ventral language pathways in persistent developmental stuttering. Cortex 2016: 81: 79–92.
Li G, Wang L, Shi F, Lyall AE, Lin W, Gilmore JH, et al. Mapping longitudinal development of local cortical gyrification in infants from birth to 2 years of age. J Neurosci 2014; 34: 4228–38.
Loucks T, Chon H, Han W. Audiovocal integration in adults who stutter. Int J Lang Commun Disord 2012; 47: 451–6.
Louko LJ, Edwards ML, Conture EG. Phonological characteristics of young stutterers and their normally fluent peers: preliminary observations. J Fluency Disord 1990; 15: 191–210.
Lu C, Chen C, Ning N, Ding G, Guo T, Peng D, et al. The neural substrates for atypical planning and execution of word production in stuttering. Exp Neurol 2009; 221: 11.
Lu C, Peng D, Chen C, Ning N, Ding G, Li K, et al. Altered effective connectivity and anomalous anatomy in the basal ganglia-thalamocortical circuit of stuttering speakers. Cortex 2010; 46: 49–67.
Luders E, Narr K, Thompson P, Woods R, Rex D, Jancke L, et al. Mapping cortical gray matter in the young adult brain: effects of gender. Neuroimage 2005; 26: 493–501.
Maguire G, Riley G, Yu B. A neurological basis of stuttering? Lancet Neurol 2002; 1: 407.
Maguire GA, Riley GD, Franklin DL, Gottschalk LA. Risperidone for the treatment of stuttering. J Clin Psychopharmacol 2000; 20: 479.
Maguire GA, Yu BP, Franklin DL. Alleviating stuttering with pharmacological interventions. Expert Opin Pharmacother 2004; 5: 1565–71.
Marsden CD, Obeso JA. The functions of the basal ganglia and the paradox of stereotaxic surgery in Parkinson's disease. Brain 1994; 117: 877–97.
Mayer M. Frog, where are you? New York, NY: Penguin Putnam Inc; 1969.
Melnick KS, Conture EG, Ohde RN. Phonological priming in picture naming of young children who stutter. J Speech Lang Hear Res 2003; 46: 1428–43.
Merchant H, Harrington DL, Meck WH. Neural basis of the perception and estimation of time. Annu Rev Neurosci 2013; 36: 313–36.
Mink JW. The basal ganglia: focused selection and inhibition of competing motor programs. Prog Neurobiol 1996; 50: 381–425.
Neef NE, Anwander A, Bütfering C, Schmidt-Samoa C, Friederici AD, Paulus W, et al. Structural connectivity of right frontal hyperactive areas scales with stuttering severity. Brain 2018; 141: 131–204.
Neumann K, Preibisch C, Euler HA, Gudenberg AWV, Lanfermann H, Gall V, et al. Cortical plasticity associated with stuttering therapy [Internet]. J Fluency Disord 2005; 30: 23–39.
Nieto-Castanon A, Ghosh S, Tourville J, Guenther F. Region of interest based analysis of functional imaging data. Neuroimage 2003; 19: 1303–16.
Nudelman HB, Herbrich KE, Hess KR, Hoyt BD, Rosenfield DB. A model of the phonatory response time of stutterers and fluent speakers to frequency-modulated tones. J Acoust Soc Am 1992; 92: 1882–8.
Paden EP, Yairi E, Ambrose NG. Early childhood stuttering II: initial status of phonological abilities. J Speech Lang Hear Res 1999; 42: 1113–24.

Qiao J, Wang Z, Zhao G, Huo Y, Herder CL, Sikora CO, et al. Functional neural circuits that underlie developmental stuttering. PLoS One 2017; 12: e0179255.

Rankin KP, Salazar A, Gorno-Tempini ML, Sollberger M, Wilson SM, Pavlic D, et al. Detecting sarcasm from paralinguistic cues: anatomic and cognitive correlates in neurodegenerative disease. Neuroimage 2009; 47: 2005–15.

Raznahan A, Shaw P, Lalonde F, Stockman M, Wallace GL, Greenstein D, et al. How does your cortex grow? J Neurosci 2011; 31: 7174–7.

Riley GD. Stuttering severity instrument: SSI-4. Austin, TX: Pro-ED Publishers; 2009.

Saltuklaroglu T, Kalinowski J, Robbins M, Crawcour S, Bowers A. Comparisons of stuttering frequency during and after speech initiation in unaltered feedback, altered auditory feedback and choral speech conditions. Int J Lang Commun Disord 2009; 44: 1000–17.

Schaer M, Ottet M-C, Scariati E, Dukes D, Franchini M, Eliez S, et al. Decreased frontal gyrification correlates with altered connectivity in children with autism. Front Hum Neurosci 2013; 7: 750.

Sommer M, Koch MA, Paulus W, Weiller C, Büchel C. Disconnection of speech-relevant brain areas in persistent developmental stuttering. Lancet 2002; 360: 380–3.

Stuart A, Frazier CL, Kalinowski J, Vos PW. The effect of frequency altered feedback on stuttering duration and type. J Speech Lang Hear Res 2008; 51: 889–97.

St. Louis KO, Hinzman AR. A descriptive study of speech, language, and hearing characteristics of school-aged stutterers. J Fluency Disord 1988; 13: 331–55.

Tourville J, Guenther F. Automatic cortical labeling system for neuroimaging studies of normal and disordered speech. Soc Neurosci 2012.

Toyomura A, Fujii T, Kuriki S. Effect of external auditory pacing on the neural activity of stuttering speakers. Neuroimage 2011; 57: 1507–16.

Usler E, Smith A, Weber C. A lag in speech motor coordination during sentence production is associated with stuttering persistence in young children. J Speech Lang Hear Res 2017; 60: 51–61.

Walsh B, Mettel, KM, Smith, A. Speech motor planning and execution deficits in early childhood stuttering. J Neurodev Disord 2015; 7: 27.

Walsh B, Tian F, Tourville JA, Yücel MA, Kuczek T, Bostian AJ. Hemodynamics of speech production: an fNIRS investigation of children who stutter. Sci Rep 2017; 7: 4034.

Watkins KE, Smith SM, Davis S, Howell P. Structural and functional abnormalities of the motor system in developmental stuttering. Brain 2008; 131: 50–9.

White T, Su S, Schmidt M, Kao C-Y, Sapiro G. The development of gyrification in childhood and adolescence. Brain Cogn 2010; 72: 36–45.

Wieland EA, McAuley JD, Dilley LC, Chang S-E. Brain and language. Brain Lang 2015; 144: 26–34.

Wolk L, Edwards ML, Conture EG. Coexistence of stuttering and disordered phonology in young children. J Speech Lang Hear Res 1993; 36: 906–17.

Yairi E, Ambrose NG, Paden EP, Throneburg RN. Predictive factors of persistence and recovery: pathways of childhood stuttering. J Commun Disord 1996; 29: 51–77.

Yairi E, Ambrose NG. Early childhood stuttering I: persistency and recovery rates. J Speech Lang Hear Res 1999; 42: 1097–12.

Yang DY-J, Beam D, Pelphrey KA, Abdullahi S, Jou RJ. Cortical morphological markers in children with autism: a structural magnetic resonance imaging study of thickness, area, volume, and gyrification. Mol Autism 2016; 7: 11.

Yaruss JS, Quesal RW. Overall Assessment of the Speaker's Experience of Stuttering (OASES): documenting multiple outcomes in stuttering treatment. J Fluency Disord 2006; 31: 90–115.

Zhang Y, Zhao L, Bi W, Wang Y, Wei G, Evans A, et al. Effects of long-term diving training on cortical gyrification. Sci Rep 2016; 6: 28243.