Individual Differences in Speech Monitoring:

Functional and Structural Correlates of Delayed Auditory Feedback

Muge Ozker[1,2], Laura Giglio[3], Ahmad Beyh[4], Stephanie J. Forkel[1,2,5], Peter Hagoort[1,2]

1. Max Planck Institute for Psycholinguistics, Netherlands

2. Donders Institute for Brain, Cognition and Behaviour, Radboud University, Netherlands

3. Institute for the Interdisciplinary Study of Language Evolution, University of Zurich, Switzerland

4. Department of Psychiatry, Brain Health Institute, Rutgers University, Piscataway, NJ, 08854, USA

5. Brain Connectivity and Behavior Laboratory, Sorbonne Universities, Paris, France

ORCHIDs: MO: 0000-0001-7472-4528, LG: 0000-0001-5369-4331, AB: 0000-0001-9071-0150, SJF: 0000-0003-0493-0283, PH: 0000-0001-7280-7549

Keywords: Speech Motor Control, Delayed Auditory Feedback (DAF), *f*MRI, DWI, Arcuate Fasciculus

Abstract

Speech production relies on continuous self-monitoring to ensure that produced sounds match intended targets. Delayed auditory feedback (DAF) disrupts speech fluency by perturbing this alignment and offers a powerful tool to study the sensorimotor control of speech. Here, we combined functional and diffusion-weighted MRI in 31 participants producing words with and without DAF. Participants showed substantial variability in how much their speech slowed under DAF, quantified by a continuous susceptibility index (SI). At the group level, DAF elicited increased activation in a right-lateralized speech motor network encompassing the superior temporal, supramarginal, inferior frontal, precentral, and supplementary motor areas, along with engagement of the left cerebellum. In contrast, individual differences revealed the opposite pattern: higher susceptibility was associated with stronger activation in left-hemisphere speech motor homologues. Diffusion analyses of the arcuate fasciculus (AF) further showed that greater fiber density in the right posterior AF predicted lower susceptibility, whereas larger right long AF volume predicted higher susceptibility. Together, these findings identify a lateralized and tract-specific organization of speech monitoring: fluency under perturbation depends on a right-lateralized monitoring network that supports integration of auditory and somatosensory feedback through the right posterior AF, while stronger auditory–motor coupling via the right long AF and excessive left-hemisphere activation predict greater vulnerability to feedback disruption.

INTRODUCTION

During speech production, the brain continuously monitors how closely the sounds being produced match the intended outcome. Hearing one's own voice is central to this monitoring process, as auditory feedback enables the detection and correction of speech errors in real time, ensuring that motor commands produce the desired speech sounds. Disrupting auditory feedback can alter speech output, as many people have experienced when speaking while listening to loud music through headphones, often raising their voice unintentionally because they cannot accurately regulate its loudness.

To investigate the neural mechanisms underlying this monitoring system, researchers often manipulate auditory feedback during speech production to introduce controlled mismatches between expected and perceived sounds. Such perturbations simulate speech errors and reveal how the brain adjusts ongoing production to maintain accuracy and fluency. Different manipulations of auditory feedback disrupt distinct aspects of speech. Perturbing the pitch (F0) of auditory feedback affects vocalization, typically causing speakers to shift their pitch in the opposite direction of the perturbation (Burnett et al., 1998; Chen et al., 2007). Altering the formant frequencies (F1, F2) disrupts articulation, prompting speakers to adjust their articulatory configurations to produce the intended vowel sound (Daliri et al., 2020; Munhall et al., 2009). Delaying auditory feedback disrupts speech coordination and fluency by introducing an asynchrony between articulation and auditory perception, often leading speakers to slow down, pause or repeat syllables to restore alignment between production and feedback (Fairbanks, 1955; Lee, 1950; Stuart et al., 2002).

Most neuroimaging studies on speech monitoring have used pitch (Behroozmand et al., 2015; Franken et al., 2018; Kort et al., 2014; Toyomura et al., 2007) or formant perturbations (Niziolek & Guenther, 2013), which mainly reveal the brain mechanisms underlying vocalization and articulation control. In contrast, despite its robust effects on

speech fluency, delayed auditory feedback (DAF) processing remains comparatively underexplored (Hashimoto & Sakai, 2003; Ozker et al., 2022, 2024; Takaso et al., 2010). A deeper understanding of DAF is particularly important because it taps into the temporal coordination processes that are fundamental to fluent speech.

In addition, individuals vary widely in their susceptibility to altered auditory feedback, yet the neural substrates underlying this variability remain poorly understood. Clinical studies highlight striking differences: DAF induces more dysfluencies in individuals with schizophrenia (Goldberg et al., 1997) or autism spectrum disorder (Lin et al., 2015) than in neurotypical controls, whereas patients with conduction aphasia show reduced susceptibility (Boller & Marcie, 1978). Intriguingly, while DAF often elicits stutter-like speech in neurotypical speakers, it can improve fluency in people who stutter (Kalinowski & Stuart, 1996). Even within neurotypical populations, substantial interindividual variability in DAF susceptibility has been reported (Agnew et al., 2018; Burke, 1975; Howell & Archer, 1984; Ozker & Hagoort, 2025), yet its neural basis remains unknown.

Current models of speech production describe speech monitoring as a dynamic interaction between perisylvian motor and sensory regions through feedforward and feedback processes (Houde & Nagarajan, 2011; Tourville & Guenther, 2011). However, our understanding of this process is shaped largely by cortical activation patterns observed in neuroimaging studies, with comparatively little knowledge on the contribution of subcortical structures and the white matter pathways that integrate the different components of the speech network. A central white matter pathway that arches around the lateral fissure, and link the frontal motor, parietal somatosensory, and temporal auditory regions relevant for speech production is the arcuate fasciculus (AF). The AF has long been recognized as critical for healthy speech and language processing (Basile et al., 2024; Catani et al., 2005; Forkel et al., 2022) and implicated in language disorders like aphasia (Dronkers, 2000; Forkel et al.,

2014, 2020) and auditory hallucinations in schizophrenia (Catani et al., 2011). Anatomically, the AF can be subdivided into three components: a direct long fronto-temporal segment, and two indirect fronto-parietal and parieto-temporal tracts that link these regions via the inferior parietal lobe (Catani et al., 2005). Different segments of the AF have been linked to distinct components of speech and language processing, including fluency (Ivanova et al., 2021), naming (Basile et al., 2024), repetition (Forkel et al., 2020), and lexical–semantic processing (Glasser & Rilling, 2008). However, their specific contributions to sensory feedback control of speech remain unclear. Notably, the AF shows marked interindividual variability in its anatomical structure (e.g. presence/robustness of the long segment, lateralization, fiber capacity)(Bernard et al., 2019; Catani et al., 2007), raising the question of whether this structural variability contributes to individual differences in susceptibility to DAF.

To address these questions, we conducted a speech production experiment in neurotypical individuals using DAF during *f*MRI, alongside diffusion-weighted imaging (DWI) and in-scanner voice recordings. Our objectives were to: (1) identify cortical networks involved in monitoring speech coordination and fluency through DAF manipulation, (2) determine the role of the AF in supporting this process, and (3) establish the neural substrates that underlie individual differences in susceptibility to DAF measured by the degree of speech slowing.

METHODS

Participants

31 native Dutch speakers (18 females, age: 24 ± 6 years) were recruited through the Radboud SONA research participation system. All participants presenting without a history of hearing or language impairments were included in the study after providing written informed consent prior to the start of the experiments. The research was approved by the

Ethics Committee of the Faculty of Social Sciences at Radboud University (ECSW-2020-046) and conducted in accordance with the Declaration of Helsinki.

Experimental Procedure

The experiment consisted of three conditions: passive listening, speaking with immediate auditory feedback (no-DAF), and speaking with delayed auditory feedback (DAF). For the passive listening condition, participants were presented with the audio recordings of 20 different words spoken by a female native Dutch speaker. The words consisted of three syllables and had a mean duration of 0.72 seconds. During passive listening trials, participants were instructed to look at the crosshair in the middle of the screen and passively listen to presented words. For the two speaking conditions, the same list of words was presented on the screen as text, one word per trial. The words appeared as red text and participants were instructed to read them out loud as soon as the text turned green. During these speaking trials, auditory feedback was presented either simultaneously (no-DAF) or with a 200 ms delay (DAF) through insert earphones. Each of the 20 words was repeated 2 times for each condition, which resulted in 40 trials per condition and a total of 120 trials for all three conditions. Trials for the three conditions were presented randomly in an event-related *f*MRI design, with randomly distributed number of rest trials in between them. During the rest trials, subjects passively fixated on the crosshair on the screen. *OptSeq* (Dale, 1999) was used for the randomization and balanced distribution of conditions and rest trials intervals across the run.

Neuroimaging Data Acquisition

Structural (T1w, diffusion-weighted) and functional (*f*MRI) data were acquired on a 3T Siemens MAGNETOM Prisma Fit scanner using a 32-channel head coil at the Donders

Institute, Nijmegen. Structural data were acquired using a T1-weighted 3D magnetization-prepared rapid gradient-echo (MPRAGE) sequence (TR = 2300 ms, TE = 3.03 ms, TI = 1100 ms, flip angle = 8°, 192 sagittal slices, voxel size = 1 × 1 × 1 mm^3, FOV = 256 × 256 mm^2, GRAPPA acceleration factor = 2). The total acquisition time was 5 minutes and 21 seconds.

Diffusion-weighted imaging was acquired as a single-shot spin-echo echo-planar imaging (EPI) acquisition with GRAPPA acceleration (factor = 2) and multi-band acceleration (factor = 3). Acquisition parameters were TR = 2940 ms, TE = 74.8 ms, flip angle = 90°, FOV = 216 × 216 mm^2, matrix = 120 × 120, isotropic voxel size of 1.8 mm^3, 81 axial slices with interleaved acquisition, and anterior-to-posterior phase encoding. Diffusion gradients were applied along 86 non-collinear directions at a b-value of 1250 s/mm^2 and 85 non-collinear directions at a b-value of 2500 s/mm^2. Twelve interleaved volumes were collected without diffusion weighting (b0 volumes). Total scan time was approximately 10 minutes. Additionally, seven b0 volumes were acquired with the reverse phase encoding direction (posterior-to-anterior) to facilitate susceptibility distortion correction.

Functional data were acquired using a T2* weighted gradient-echo echo-planar imaging (EPI) sequence with a sparse sampling design (TR = 3000 ms, TA ≈ 1000 ms, TE = 33 ms, flip angle = 80°). Each volume consisted of 52 interleaved transversal slices (voxel size = 2.4 × 2.4 × 2.4 mm^3, FOV = 210 × 210 mm^2) acquired with a multiband acceleration factor of 4. To minimize contamination of auditory stimuli by scanner noise, a sparse sampling design was used: image acquisition lasted ~1 s, followed by a 2 s silent gap during which stimuli were presented. A total of 400 volumes were collected per run. All functional data were collected in a single acquisition lasting approximately 20 minutes. Stimulus presentation was executed using Matlab Psychtoolbox-3 running on an HP Intel Core i7 laptop with a Windows 10 operating system. As participants read aloud the words, their voices were recorded using an MRI compatible microphone (Optoacoustics, FOMRI-III) and

played back using a Focusrite Scarlett 2i2 audio interface and S14 MRI compatible insert earphone. For the DAF condition, the recorded audio was delayed using Matlab's Audio Toolbox functions.

DWI Analysis and Tractography

The preprocessing of the DWI data followed our in-house pipeline (Beyh et al., 2025; Ohlerth et al., 2025). Briefly, DWI data were corrected for thermal noise (Veraart et al., 2016) and Gibbs ringing artefacts (Kellner et al., 2016) using tools implemented in *TORTOISE* (Irfanoglu et al., 2017). The pair of b0 images acquired with reverse phase encoding was used to estimate a geometric distortion field in *topup* (Andersson et al., 2003). The tool *eddy* was then used to correct the full DWI series for head motion, eddy current distortions, slice-to-volume motion, signal dropout, and susceptibility distortions based on the *topup* field, including the effect of motion on these distortions (Andersson et al., 2016, 2017, 2018; Andersson & Sotiropoulos, 2016). DWI data from the b = 2500 s/mm^2 shell were used for tractography. The data were modelled in StarTrack (https://nbl-research.github.io) using the damped Richardson-Lucy spherical deconvolution algorithm (Dell'acqua et al., 2010; Dell'Acqua, Simmons, et al., 2013) and the following parameters: fiber response function α = 1.5; number of iterations = 200; damping parameter η = 0.0015; adaptive regularization parameter v = 16. Tractography was then performed using a multi-fiber Euler-like algorithm and the following parameters: minimum amplitude = 0.0025; angle threshold = 40°; step size = 0.9 mm; fiber length range = 20–300 mm. The anisotropic power (AP) maps (Dell'Acqua, Lacerda, et al., 2013) generated by StarTrack were used to normalize each participant's DWI data and tractography to the common space of the ICBM 2009a nonlinear symmetric MNI template (Fonov et al., 2011) using diffeomorphic registration in ANTs (Avants et al., 2008, 2014). MegaTrack (Dell'Acqua et al., 2025) was used to perform semi-automatic virtual

dissections using the region-of-interest placement described in Forkel et al. (Forkel et al., 2025). The use of the symmetric template allowed us to use the symmetric flipping option in MegaTrack, which renders the operator blind to the hemisphere they are dissecting, hence removing any user bias. The interactive virtual dissections were performed in TrackVis (http://trackvis.org). Once the dissections were complete, tract volume and the hindrance modulated orientational anisotropy (HMOA) (Dell'Acqua, Simmons, et al., 2013) were extracted in native space for each participant and each tract. Tract volume (in milliliters, mL) was calculated by multiplying the volume of a single voxel (5.832×10-3 mL) by the total number of voxels visited by the tract. HMOA was extracted directly from the tractography data at each streamline point and was averaged over the entire tract. HMOA corresponds to the amplitude of the fiber orientation distribution (unitless) at each point in the tractogram and reflects the apparent fiber density at that point, among other measures of microstructural organization. For the main analysis, this resulted in three tracts (the three AF segments) in each hemisphere, with two metrics each (volume for macrostructure and HMOA for microstructure), resulting in a total of 12 variables that were tested for correlations with SI. The false discovery rate (FDR) method (Benjamini & Hochberg, 1995) was used to compensate for these multiple comparisons.

*f*MRI Analysis

Preprocessing: *f*MRI preprocessing was performed using fMRIPrep 23.1.0 (Esteban et al., 2019; Markiewicz et al., 2025). The T1-weighted images were spatially normalized to MNI standard space (MNI152NLin2009cAsym) after skull stripping and segmentation. The functional images were skull stripped, head motion parameters were estimated, followed by slice-time correction. The BOLD time-series were then co-registered to the T1w image and

finally resampled to standard space (see *Supplementary Material* for a detailed description of the fMRIPrep preprocessing pipeline).

General Linear Model (GLM) analysis: First, we smoothed the preprocessed functional images in MNI space with an isotropic Gaussian kernel of 4 mm FWHM in SPM12 in Matlab R2023b. We then computed a general linear model in SPM12 with the following condition regressors: DAF, no-DAF, and Listen, including their temporal derivatives. In addition, we added head motion parameters estimated by fMRIPrep as confound regressors. The model was estimated after high-pass filtering and then we computed contrasts between the DAF and no-DAF conditions using weights [1 -1]. The contrast images were entered into a group-level one-sample t-test, where we also included participant-specific susceptibility indices (see *Behavioural Analysis and Susceptibility Index (SI) Calculation* section) as a covariate. We then generated contrasts for the second-level model for the DAF and no-DAF difference and for the SI covariate. Statistical maps were thresholded at the voxel level ($p < 0.001$ uncorrected) and subjected to a cluster-level extent threshold ($p < 0.05$ family-wise error corrected at the cluster-level) to identify significant activation clusters. Visualization was done in Surf Ice (https://www.nitrc.org/projects/surfice) and MRIcron (https://www.nitrc.org/projects/mricron).

Behavioural Analysis and Susceptibility Index (SI) Calculation

Participants' audio recordings were analyzed in *Praat* (Boersma, 2001). For each trial, word onset and offset were manually marked on the audio waveform to determine word duration. A paired t-test was performed to test the effect of auditory feedback delay on word duration, comparing the DAF and no-DAF conditions across participants. To quantify individual susceptibility to DAF, a susceptibility index (SI) was computed for each participant comparing word durations for the different conditions using the formula:

$$SI = (DAF - no\ DAF) / (DAF + no\ DAF).$$

This index ranged from −1 to 1, with higher values indicating greater susceptibility to DAF. One participant was excluded from the SI analysis because the participant's voice could not be recorded during *f*MRI scanning.

RESULTS

Behavioural Results

For the behavioral analysis, we examined participants' voice recordings acquired during scanning by calculating word duration for each trial. Speaking with 200 ms delayed auditory feedback (DAF) slowed speech rate compared to speaking with immediate feedback (no-DAF), as reflected by a significant increase in word duration across participants (no-DAF: 0.85 ± 0.02 s; DAF: 0.95 ± 0.03 s; paired t-test: $t = 7.859$, $p < 0.001$; **Figure 1A**). Individual inspection of participants revealed substantial variability in susceptibility to DAF: some participants showed pronounced slowing, whereas others largely maintained their speech rate. To quantify this effect, a susceptibility index (SI) was computed for each participant by comparing word duration between the DAF and no-DAF conditions (see *Methods* for SI calculation). SI values varied widely across participants, ranging from 0.011 to 0.131. (Participant #1 versus #18 in **Figure 1B**).

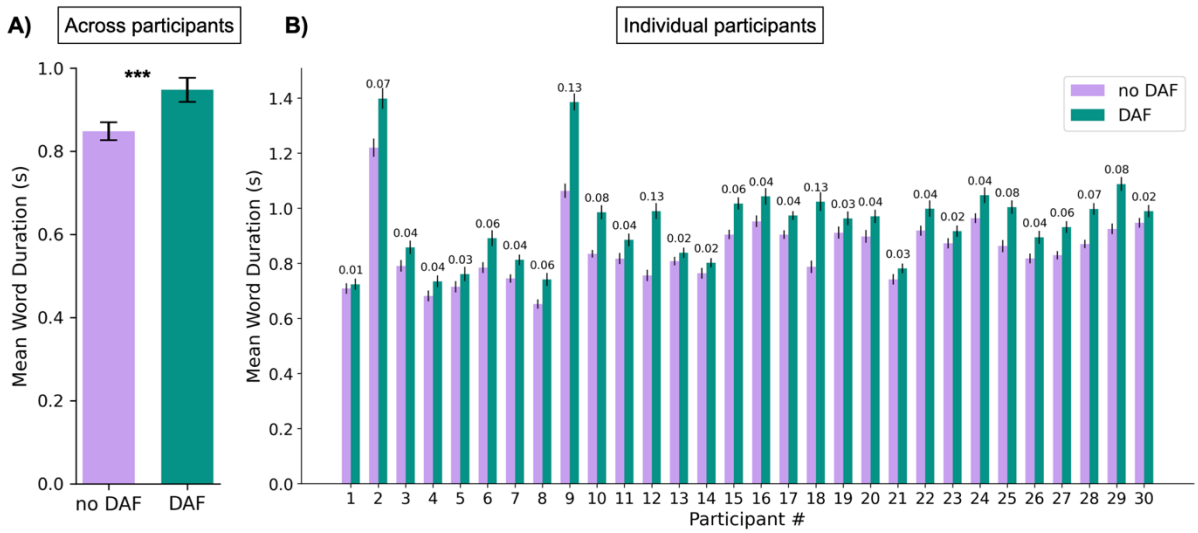

Figure 1: Delayed auditory feedback (DAF) slows speech. (A) *Group mean (± SD) word duration during reading with immediate feedback (no-DAF) versus 200ms DAF. DAF significantly lengthened word duration (paired t-test: $t = 7.859$, $p < 0.001$).* ***(B)*** *Word duration for each participant in both conditions. Numbers above bars show the DAF susceptibility index (SI). Larger SI indicates greater slowing under DAF (SI shown for participants with usable audio).*

Group-Level Functional Activation for DAF

Across the whole brain, producing speech under delayed auditory feedback (DAF > no-DAF) elicited stronger and more widespread activation in right-hemisphere speech motor regions, indicating greater recruitment of this network for speech monitoring when auditory feedback is disrupted (**Table S1, Figure 2A–B**). Significantly larger activation was observed in the bilateral superior temporal gyri (STG) extending into the supramarginal gyrus (SMG), consistent with increased auditory and somatosensory processing demands. Enhanced responses were also found in speech motor regions, including the right inferior frontal gyrus (IFG), right supplementary motor area (SMA), and bilateral precentral gyri (PreCG), reflecting heightened demands on motor planning and articulatory control under perturbed feedback conditions.

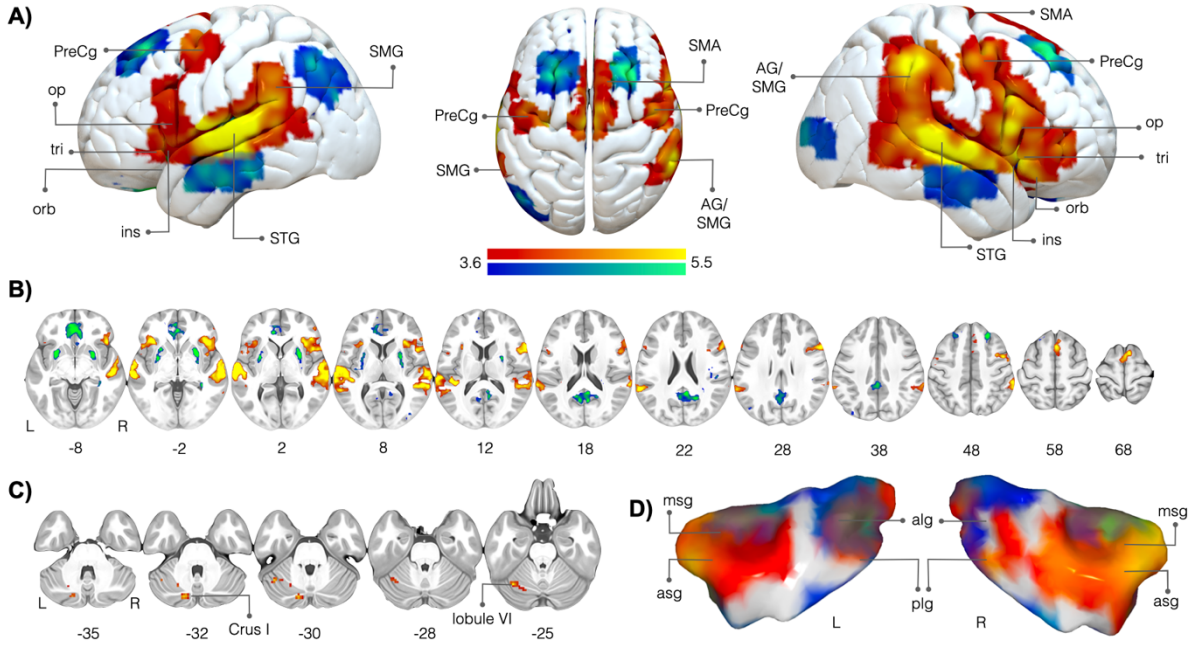

Figure 2: Whole-brain activations during speech production with delayed auditory feedback (DAF).
(A) *Cortical surface renderings (MNI152) showing group-level activation for the contrast* DAF > no-DAF *(warm colors) and the reverse contrast* no-DAF > DAF *(cool colors), displayed on left lateral, superior, and right lateral views.* ***(B)*** *Axial sections showing the same effects across the whole brain* ***(C)*** *Cerebellar sections* ***(D)*** *Left and right insular surfaces (Faillenot et al., 2017).*
SMA, supplementary motor area; PreCg, precentral gyrus; STG, superior temporal gyrus; SMG/AG, supramarginal gyrus/angular gyrus; IFG op/tri/orb, inferior frontal gyrus pars opercularis/triangularis/orbitalis; ins, insula. asg (anterior short gyrus), msg (middle short gyrus), alg (anterior long gyrus), plg (posterior long gyrus).

Additional activation was detected in the left superior cerebellar regions, including Crus I and lobule VI (**Figure 2C**), which have been previously shown to engage during both overt and covert speech production and are considered integral to the online control of articulatory movement sequences (Ackermann & Riecker, 2010). Significant activation was also observed in the insula, most prominently in the anterior portion of the right insular gyrus (anterior and middle short gyri) (**Figure 2D**). This portion of the insula has been consistently implicated in overt speech production (Kurth et al., 2010), and damage to it is known to cause impairments in speech and orofacial motor control (Dronkers, 1996; Marien et al., 2001).

The reverse contrast (no-DAF > DAF) revealed regions more active when speech unfolds with natural, unperturbed feedback. This analysis identified a frontal–parietal/medial–temporal network encompassing the left anterior cingulate cortex (ACC),

right precuneus/posterior medial cortex, left angular gyrus, and bilateral hippocampal formation (**Table S1, Figure 2A–B**). This activation pattern overlaps substantially with the canonical default mode network (Raichle et al., 2001), suggesting that when speech proceeds with intact feedback, reduced demands for external monitoring allow internally oriented processes to remain more active. Increased activity was also observed in the bilateral putamen. Given that the putamen is associated with motor sequence learning (Janacsek et al., 2020) and strongly responds to successful movement outcomes (Brown et al., 2006; Seidler et al., 2005), its increased activation for the no-DAF condition likely reflects the reinforcement of successfully produced articulatory sequences when speech outcomes align with internal predictions. Finally, smaller clusters (< 75 voxels) of activation were identified in the right Rolandic operculum, bilateral middle temporal cortices, left fusiform gyrus, right middle occipital gyrus, and right insula.

Functional Associates of DAF Susceptibility

To investigate how brain activation patterns varied with individual susceptibility to DAF, we incorporated susceptibility indices (SI) as covariates in the group-level GLM analysis. This approach revealed a predominantly left-lateralized activation pattern (**Table S2, Figure 3A–B**). Higher DAF susceptibility was associated with larger responses in the left inferior frontal gyrus (IFG), dorsal precentral gyrus (PreCG), supramarginal gyrus (SMG), anterior cingulate cortex (ACC), and middle frontal gyrus (MFG), accompanied by strong engagement of the cerebellar vermis (lobules 4–5) (**Figure 3C**). Significant clusters were also observed in the right hemisphere, including the supplementary motor area (SMA), IFG, and cerebellar lobule VI. In both hemispheres, IFG activation extended into the anterior insula (anterior and middle short gyri), with particularly broad activation in the left insular region

(**Figure 3D**). For the reverse contrast (no-DAF > DAF), SI values were positively associated with activity in the left angular gyrus.

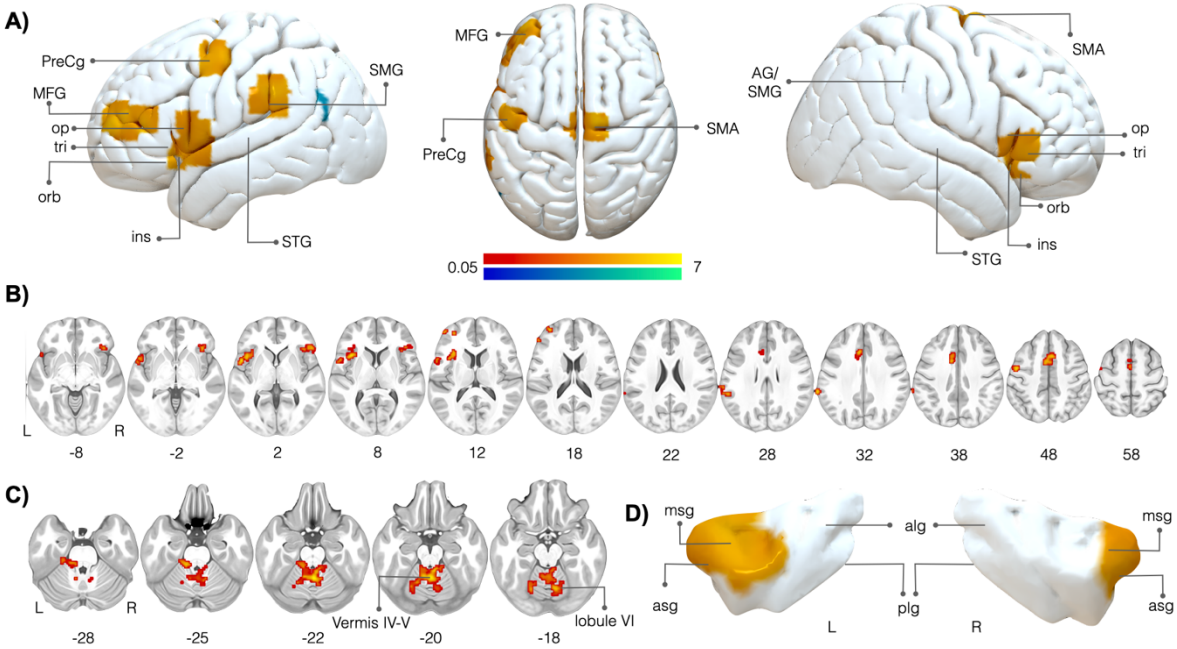

Figure 3: Whole-brain activations associated with individual susceptibility to delayed auditory feedback (DAF).
(A) *Cortical surface renderings (MNI152) showing regions where activation in the contrast* DAF > no-DAF *(warm colors) and the reverse contrast* no-DAF > DAF *(cool colors) covaried with participants' DAF susceptibility indices (SI), displayed on left lateral, superior, and right lateral views.* ***(B)*** *Axial sections showing the same effects across the whole brain* ***(C)*** *Cerebellar sections* ***(D)*** *Left and right insular surfaces (Faillenot et al., 2017).*
SMA, supplementary motor area; PreCg, precentral gyrus; STG, superior temporal gyrus; SMG/AG, supramarginal gyrus/angular gyrus; IFG op/tri/orb, inferior frontal gyrus pars opercularis/triangularis/orbitalis; MFG, middle frontal gyrus; ins, insula. asg (anterior short gyrus), msg (middle short gyrus), alg (anterior long gyrus), plg (posterior long gyrus).

Overall, the activation patterns that scale with individual susceptibility exhibit a lateralization opposite to the average group-level effect observed for DAF > no-DAF, which showed stronger right-hemisphere activity (**Figure 2**). These results indicate that participants who were more behaviorally affected by DAF show greater activation in left hemisphere speech motor regions, along with cerebellar areas.

Structural Associates of DAF Susceptibility

Motivated by the perisylvian *f*MRI activation patterns in response to DAF (**Figure 2A-B**) and the putative role of arcuate pathways in sensory-motor integration (Catani et al., 2005), we tested whether interindividual anatomical variation in the arcuate fasciculus (AF) segments relates to DAF susceptibility. We used subject-specific tractography to reconstruct the anterior (frontal–parietal), posterior (temporal–parietal), and long (temporal–frontal) arcuate segments in each hemisphere (**Figure 4A-B**). All three segments were dissected in both hemispheres for every participant. For each tract, volume and hindrance-modulated orientational anisotropy (HMOA) (Dell'Acqua, Simmons, et al., 2013) were extracted and correlated with individual susceptibility indices.

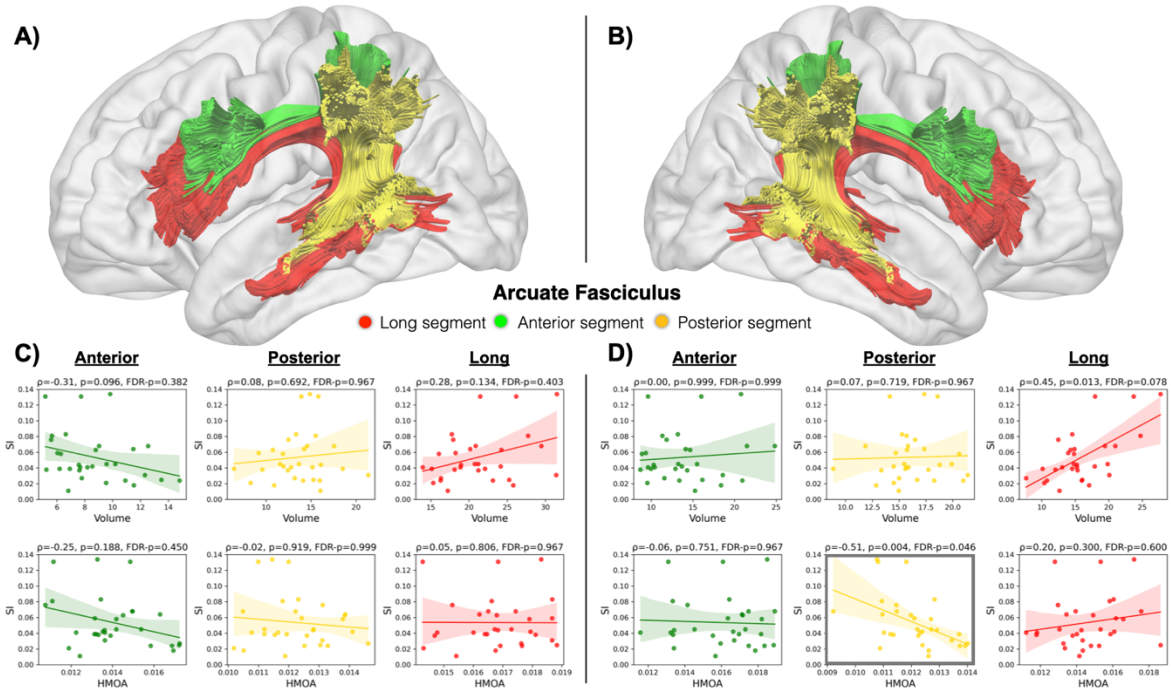

Figure 4: ***Correlations between participants' DAF susceptibility and the structural integrity of the three arcuate fasciculus segments.***
(A-B) *Tractography-based dissection of the arcuate fasciculus (AF) into three segments: anterior (fronto-parietal, posterior (parietal-temporal), and long (fronto-temporal)* ***(C)*** *Left hemisphere: Associations (Spearman correlations) between SI and AF segment volume and HMOA; points are participants, lines show fitted trend with 95% CI* ***(D)*** *Right hemisphere: matching analyses for the homologous segments. Significant negative correlation was found for right posterior segment (scatter plot is shown in a darker gray frame). Spearman rho, p-value and FDR corrected p-values are shown on top of the plots.*
HMOA, hindrance-modulated orientational anisotropy; SI: Susceptibility index.

DAF susceptibility was negatively associated with HMOA of the right posterior segment (Spearman correlation: $p = 0.004$, FDR-$p = 0.046$), indicating that greater fiber density between posterior temporal and inferior parietal territories was linked to a reduced behavioral impact of DAF. A positive association was observed between SI and volume of the right long segment ($p = 0.013$), which did not survive FDR correction (FDR-$p = 0.078$). No significant associations were observed for the right anterior segment or for any left hemisphere AF segments (**Figure 4C-D**). These results point specifically to the right temporal–parietal arcuate as a structural support against DAF-induced dysfluency. For completeness, we also surveyed other canonical language-relevant white-matter pathways (e.g., frontal aslant tract and frontal-insular tracts), but none of these tract–behavior correlations survived multiple comparison correction (**Figure S1 and S2**).

DISCUSSION

Fluent speech depends on the alignment between internally predicted and externally perceived sensory feedback. When this alignment is perturbed, as under delayed auditory feedback (DAF), the brain must reconcile conflicting sensory inputs to maintain fluency. By linking behavioral susceptibility with both functional activation and white-matter connectivity, our results show that individual differences in this compensatory process reflect underlying neural architecture: resilience to DAF engages a right-lateralized network supporting integration of auditory and somatosensory feedback via anatomical connections, whereas vulnerability is characterized by stronger structural auditory–motor coupling and increased activation of left-hemisphere speech monitoring circuits.

Speech production models commonly describe sensorimotor interactions as a feedforward-feedback control system (Houde & Nagarajan, 2011; Tourville & Guenther,

2011). Within this framework, when a motor command is issued, an efference copy generates a neural prediction of the expected sensory consequences. This internal prediction is compared with incoming sensory input, enabling online correction of any discrepancies. Mismatch errors occur when the perceived sensory feedback diverges from the efference-based prediction. Such mismatches are detected by auditory comparator mechanisms in the posterior superior temporal gyrus (pSTG), as well as somatosensory comparators in the supramarginal gyrus (SMG). Corrective signals are transmitted to the inferior frontal gyrus (IFG), where articulatory plans are adjusted in real-time and relayed to the motor cortex, which subsequently sends articulatory commands to the vocal tract. This parallel system operating for auditory and somatosensory feedback (e.g., tactile sensations from the mouth and proprioception from the articulators) is thought to create a dual-control mechanism for speech production.

Our functional activation results converge on these described nodes of the feedforward-feedback loop. Consistent with a heightened demand on sensory processing and motor control in response to altered auditory feedback, DAF elicited activation in the STG, SMG, IFG and precentral gyrus (PreCg) (**Figure 2**), predominantly in the right hemisphere. Specifically, the dorsal PreCg activation aligns with our earlier electrocorticography findings in epilepsy patients, which showed selective engagement of this region during sentence production under DAF, supporting its role in maintaining speech fluency when dynamic auditory feedback processing is required to produce syllable sequences (Ozker et al., 2022).

Beyond these core regions, the right hemisphere showed significant activation in the anterior portion of the insular gyrus, which is structurally connected with the IFG via frontal-insular tracts (Catani et al., 2012; Quabs et al., 2025; Rojkova et al., 2016). Consistent with its proposed role as a relay between speech planning in IFG and motor execution in

subcortical and cortical speech circuits (Eickhoff et al., 2009), anterior insula activation may reflect the translation of phonetic intentions into coordinated vocal tract movements. Similarly, the SMG is thought to act as a key relay in the dorsal speech stream, supporting sensory-motor transformations by mapping phonetic representations onto articulatory movements (Hickok & Poeppel, 2007; Rauschecker & Scott, 2009). These activations likely represent a specialized network for integrating phonetic planning with sensorimotor control, enabling the rapid error monitoring and subsequent motor adjustments under DAF.

We also observed engagement of the right supplementary motor area (SMA) together with left cerebellar regions, which are structurally connected via crossed cortico–ponto–cerebellar and cerebello–thalamo–cortical pathways (Akkal et al., 2007). Moreover, the SMA is reciprocally connected with the basal ganglia via thalamocortical loops and linked to the IFG via the frontal aslant tract (Thiebaut de Schotten et al., 2012), forming a circuit that supports the initiation, sequencing, and timing of speech motor commands (Bullock et al., 2025; Hertrich et al., 2016). Consistent with these crossed pathways, co-activation of the right SMA with left cerebellar structures during DAF suggests increased demands on the temporal precision and coordination of articulatory movements.

In contrast to the right-hemisphere bias observed at the group level in response to DAF, incorporating individual DAF susceptibility indices revealed a distinct shift in activation patterns (**Figure 3**). Participants who were more susceptible to DAF showed stronger activation in several left-hemisphere speech motor homologues, including the IFG, SMG, and PreCG. This extension of cortical engagement from a right-dominant speech motor control network toward a left-weighted pattern suggests that individuals more affected by feedback disruption exhibit broader and poorly inhibited activation across left hemisphere speech motor regions, likely reflecting overactive monitoring processes. In contrast, more

resilient individuals appear to engage a more focused, right-lateralized monitoring network that supports efficient speech motor control under altered feedback. In addition, greater DAF susceptibility was associated with heightened activity also in the ACC and MFG, the latter corresponding to the dorsolateral prefrontal cortex (DLPFC). Activation of these higher-level conflict monitoring and executive control networks (Botvinick et al., 2001; Carter et al., 1998; Miller & Cohen, 2001) further supports the interpretation of excessive monitoring in highly susceptible individuals.

Given that speech monitoring relies on rapid information exchange across auditory, somatosensory and motor speech regions, we next asked whether interindividual differences in the white matter linking them could further explain variability in DAF susceptibility. We reconstructed the arcuate fasciculus in each hemisphere and examined its three components (**Figure 4**): the long temporo-frontal segment, the posterior parieto-temporal segment, and the anterior fronto-parietal segment (Catani et al., 2005), whose cortical terminations overlapped with our functional cluster peaks in IFG, SMG and pSTG (Friederici, 2009). Across participants, susceptibility to DAF increased with the volume of the right long segment (uncorrected p = 0.013, FDR-p = 0.078). Although this trend did not survive multiple comparison correction, it is anatomically coherent (**Figure 5**): a more substantial right temporo-frontal

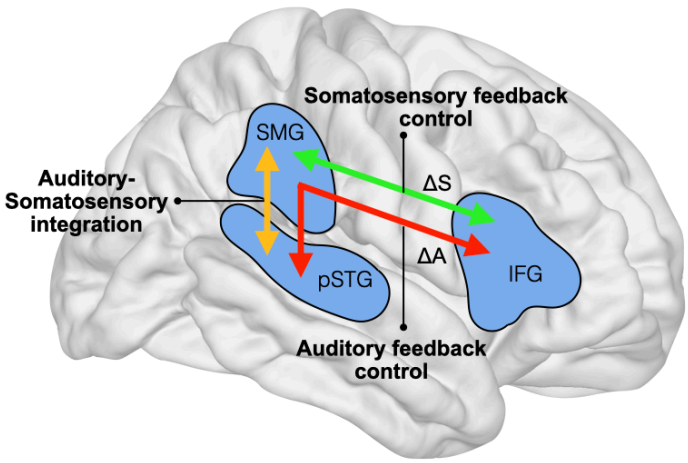

Figure 5: Illustration of sensory feedback control during speech production. *The posterior superior temporal gyrus (pSTG) and supramarginal gyrus (SMG) serve as key nodes for encoding auditory and somatosensory corrective signals (ΔA and ΔS), respectively. These signals are then relayed to speech motor planning regions like the inferior frontal gyrus (IFG) to adjust the ongoing motor commands. The crosstalk between the pSTG and SMG enables an efficient weighing and integration of the two sensory feedback streams for motor control. The functional connections that enable this monitoring process are supported by the structural pathways of the arcuate fasciculus (AF) depicted by the double-sided arrows highlighting anterior (green), posterior (yellow) and long (red) segments of the AF.*

conduit would be expected to up-weight auditory input to frontal speech motor regions; when feedback is delayed, that heavier weighting manifests as greater behavioral slowing. This implies that individuals who are more affected by DAF may rely more heavily on auditory input, making them more vulnerable when auditory feedback is disrupted. Conversely, a significant negative correlation emerged between DAF susceptibility and fiber density (HMOA) of the right posterior AF segment (uncorrected $p = 0.004$; FDR-$p = 0.046$). This finding suggests that stronger connectivity between pSTG and SMG supports cross-sensory interactions and more effective weighting and integration of multiple feedback sources, enabling individuals to better maintain fluency when auditory and somatosensory channels provide conflicting information. DAF fundamentally disrupts sensorimotor processing by introducing a temporal misalignment between the motor commands and their auditory consequences, thereby desynchronizing auditory and somatosensory feedback streams. When the brain attempts to correct for a perturbation in auditory feedback, it may require deviating from its somatosensory targets, and vice versa. This inherent trade-off suggests that effective crosstalk and dynamic integration between auditory and somatosensory feedback systems are critical for maintaining speech fluency under feedback perturbation.

Taken together, these observations motivate a tract-specific account of feedback control, one in which the right long AF segment can amplify auditory-to-motor drive, whereas the right posterior AF segment provides a stabilizing substrate by supporting auditory–somatosensory integration. Notably, these structural effects, together with the right-lateralized functional activations under DAF, build on current speech models by highlighting a broader role of the right hemisphere in processing not only vocal pitch or formant manipulations relevant to vocalization and articulation, but also speech coordination and fluency.

In summary, by leveraging a fluency-perturbing manipulation (DAF), an individual behavioral susceptibility index, whole-brain *f*MRI, and tract-specific diffusion metrics, we provide converging functional and structural evidence that resilience or vulnerability to feedback disruption reflects differential recruitment of speech motor integration networks and variability in the strength of auditory–somatosensory connections. Our results reveal a division of labor: fluent speech under DAF relies on a right-lateralized feedback monitoring network coupled with the left cerebellar hemisphere, whereas greater individual vulnerability is marked by increased activation of left speech motor circuitry. At the structural level, a higher fiber density of the right posterior arcuate fasciculus predicts lower susceptibility, underscoring its role in supporting efficient sensory feedback integration. These findings move the field beyond group averages to a mechanistic, person-specific account of speech monitoring, with translational implications for understanding fluency breakdowns in disorders such as stuttering and relevant speech technologies.

LIMITATIONS

Despite these compelling findings, several features of our study design qualify our inferences and point to directions for future research. First, our focus on auditory perturbation precludes direct assessment of how auditory and somatosensory feedback are jointly weighted. Incorporating a somatosensory perturbation paradigm would enable a comprehensive investigation of this interaction (Tremblay et al., 2003; Weerathunge et al., 2022), while also allowing us to assess participants' somatosensory susceptibilities. Second, our behavioral index was based primarily on word duration. While this provided a continuous measure of susceptibility, richer fluency measures (e.g., pause structure, speech rate dynamics, error profiles, and articulatory kinematics) would provide a more complete behavioral characterization of DAF susceptibility. The use of a word production task, rather

than longer speech segments (e.g., sentences or paragraphs), was a necessary design choice to mitigate the high susceptibility of overt speech in the scanner to motion and susceptibility artifacts, which can reduce sensitivity in regions such as the inferior frontal, anterior temporal, and cerebellar cortices. Finally, we acknowledge the inherent limitations of our imaging modalities. Our *f*MRI results are correlational and do not establish the directionality of information flow within the speech motor network. Model-based approaches (e.g., Dynamic Causal Modeling), brain perturbation (e.g., TMS), or intracranial recordings, would provide stronger causal leverage. Additionally, while diffusion tractography is affected by crossing fibers and partial-volume effects, we used a high diffusion weight and advanced spherical deconvolution to overcome these limitations to the best of our ability. Our a priori focus on the arcuate fasciculus increases the interpretability of our findings but risks overlooking other relevant pathways. Although exploratory analyses of other tracts (FAT and FIT 1-5, **Figure S1 and S2**) did not survive multiple comparison correction, they suggest clear avenues for future hypothesis-generating work. These constraints outline a clear roadmap for the field: larger datasets with richer behavioral assays, the incorporation of somatosensory perturbations, causal and computational assays of network dynamics, and multimodal microstructure to further refine the structure–function account of feedback control.

Acknowledgements

The authors thank José Marques for programming the MRI acquisition protocol and Marteen van den Heuvel for programming the stimulus presentation. We also thank the members of the Neurobiology of Language department at the Max Planck Institute for Psycholinguistics and the Clinical Neuroanatomy of Language research group at the Donders Institute for their valuable feedback throughout this project.

Data Availability

Preprocessed fMRI data will be made available via the Radboud Data Repository (https://data.ru.nl/). Group-level statistical T-maps and correlational spreadsheets linking individual tract metrics to susceptibility indices will be shared on GitHub (https://github.com/MugeOzker).

Funding

This work was supported by the Dutch Research Council NWO Language in Interaction Grant (024.001.006), Donders Mohrmann Fellowship on 'Neurovariability' No. 2401515 (SJF), the Dutch Research Council NWO Aspasia Grant 'Human individuality: phenotypes, cognition, and brain disorders' (SJF) and the NWO SSH Open Competition grant (SJF).

REFERENCES

Ackermann, H., & Riecker, A. (2010). Cerebral control of motor aspects of speech production: Neurophysiological and functional imaging data. In B. Maassen & P. van Lieshout (Eds.), *Speech Motor Control: New developments in basic and applied research* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199235797.003.0007

Agnew, Z. K., McGettigan, C., Banks, B., & Scott, S. K. (2018). Group and individual variability in speech production networks during delayed auditory feedback. *The Journal of the Acoustical Society of America*, *143*(5), 3009. https://doi.org/10.1121/1.5026500

Akkal, D., Dum, R. P., & Strick, P. L. (2007). Supplementary motor area and presupplementary motor area: Targets of basal ganglia and cerebellar output. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(40), 10659–10673. https://doi.org/10.1523/JNEUROSCI.3134-07.2007

Andersson, J. L. R., Graham, M. S., Drobnjak, I., Zhang, H., & Campbell, J. (2018). Susceptibility-induced distortion that varies due to motion: Correction in diffusion MR without acquiring additional data. *NeuroImage*, *171*, 277–295. https://doi.org/10.1016/j.neuroimage.2017.12.040

Andersson, J. L. R., Graham, M. S., Drobnjak, I., Zhang, H., Filippini, N., & Bastiani, M. (2017). Towards a comprehensive framework for movement and distortion correction of diffusion MR images: Within volume movement. *NeuroImage*, *152*, 450–466. https://doi.org/10.1016/j.neuroimage.2017.02.085

Andersson, J. L. R., Graham, M. S., Zsoldos, E., & Sotiropoulos, S. N. (2016). Incorporating outlier detection and replacement into a non-parametric framework for movement and

distortion correction of diffusion MR images. *NeuroImage*, *141*, 556–572. https://doi.org/10.1016/j.neuroimage.2016.06.058

Andersson, J. L. R., Skare, S., & Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. *NeuroImage*, *20*(2), 870–888. https://doi.org/10.1016/S1053-8119(03)00336-7

Andersson, J. L. R., & Sotiropoulos, S. N. (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage*, *125*, 1063–1078. https://doi.org/10.1016/j.neuroimage.2015.10.019

Basile, G. A., Nozais, V., Quartarone, A., Giustiniani, A., Ielo, A., Cerasa, A., Milardi, D., Abdallah, M., Thiebaut de Schotten, M., Forkel, S. J., & Cacciola, A. (2024). Functional anatomy and topographical organization of the frontotemporal arcuate fasciculus. *Communications Biology*, *7*(1), 1655. https://doi.org/10.1038/s42003-024-07274-3

Behroozmand, R., Shebek, R., Hansen, D. R., Oya, H., Robin, D. A., Howard, M. A., & Greenlee, J. D. W. (2015). Sensory-motor networks involved in speech production and motor control: An fMRI study. *NeuroImage*, *109*, 418–428. https://doi.org/10.1016/j.neuroimage.2015.01.040

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, F., Zemmoura, I., Ter Minassian, A., Lemée, J.-M., & Menei, P. (2019). Anatomical variability of the arcuate fasciculus: A systematical review. *Surgical and Radiologic Anatomy: SRA*, *41*(8), 889–900. https://doi.org/10.1007/s00276-019-02244-5

Beyh, A., Ohlerth, A.-K., & Forkel, S. J. (2025). Harnessing Advanced Tractography in Neurosurgical Practice. In S. M. Krieg & T. Picht (Eds.), *Navigated Transcranial Magnetic Stimulation in Neurosurgery* (pp. 385–411). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-97155-6_21

Boersma, P. (2001). *Praat, a system for doing phonetics by computer.* https://www.fon.hum.uva.nl/praat/

Boller, F., & Marcie, P. (1978). Possible role of abnormal auditory feedback in conduction aphasia. *Neuropsychologia*, *16*(4), 521–524. https://doi.org/10.1016/0028-3932(78)90078-7

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652. https://doi.org/10.1037/0033-295x.108.3.624

Brown, P., Chen, C. C., Wang, S., Kühn, A. A., Doyle, L., Yarrow, K., Nuttin, B., Stein, J., & Aziz, T. (2006). Involvement of human basal ganglia in offline feedback control of voluntary movement. *Current Biology: CB*, *16*(21), 2129–2134. https://doi.org/10.1016/j.cub.2006.08.088

Bullock, L., Forseth, K. J., Woolnough, O., Rollo, P. S., & Tandon, N. (2025). Supplementary motor area in speech initiation: A large-scale intracranial EEG evaluation of stereotyped word articulation. *iScience*, *28*(1), 111531. https://doi.org/10.1016/j.isci.2024.111531

Burke, B. D. (1975). Susceptibility to delayed auditory feedback and dependence on auditory or oral sensory feedback. *Journal of Communication Disorders*, *8*(1), 75–96. https://doi.org/10.1016/0021-9924(75)90028-3

Burnett, T. A., Freedland, M. B., Larson, C. R., & Hain, T. C. (1998). Voice F0 responses to manipulations in pitch feedback. *The Journal of the Acoustical Society of America*, *103*(6), 3153–3161. https://doi.org/10.1121/1.423073

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science (New York, N.Y.)*, *280*(5364), 747–749. https://doi.org/10.1126/science.280.5364.747

Catani, M., Allin, M. P. G., Husain, M., Pugliese, L., Mesulam, M. M., Murray, R. M., & Jones, D. K. (2007). Symmetries in human brain language pathways correlate with verbal recall. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(43), 17163–17168. https://doi.org/10.1073/pnas.0702116104

Catani, M., Craig, M. C., Forkel, S. J., Kanaan, R., Picchioni, M., Toulopoulou, T., Shergill, S., Williams, S., Murphy, D. G., & McGuire, P. (2011). Altered integrity of perisylvian language pathways in schizophrenia: Relationship to auditory hallucinations. *Biological Psychiatry*, *70*(12), 1143–1150. https://doi.org/10.1016/j.biopsych.2011.06.013

Catani, M., Dell'Acqua, F., Vergani, F., Malik, F., Hodge, H., Roy, P., Valabregue, R., & Thiebaut de Schotten, M. (2012). Short frontal lobe connections of the human brain. *Cortex*, *48*(2), 273–291. https://doi.org/10.1016/j.cortex.2011.12.001

Catani, M., Jones, D. K., & ffytche, D. H. (2005). Perisylvian language networks of the human brain. *Annals of Neurology*, *57*(1), 8–16. https://doi.org/10.1002/ana.20319

Chen, S. H., Liu, H., Xu, Y., & Larson, C. R. (2007). Voice F0 responses to pitch-shifted voice feedback during English speech. *The Journal of the Acoustical Society of America*, *121*(2), 1157–1163. https://doi.org/10.1121/1.2404624

Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Human Brain Mapping*, *8*(2–3), 109–114. https://doi.org/10.1002/(SICI)1097-0193(1999)8:2/3%253C109::AID-HBM7%253E3.0.CO;2-W

Daliri, A., Chao, S.-C., & Fitzgerald, L. C. (2020). Compensatory Responses to Formant Perturbations Proportionally Decrease as Perturbations Increase. *Journal of Speech, Language, and Hearing Research: JSLHR*, *63*(10), 3392–3407. https://doi.org/10.1044/2020_JSLHR-19-00422

Dell'Acqua, F., Beyh, A., Stones, R., Barrett, R. L., Requejo, F. D. S., Laguna, P. L., Lacerda, L. M., Davison, C., Leslie, A., Howells, H., Goldstein, L. H., Williams, S. C., & Catani, M. (2025). *MegaTrack: A framework for the anatomically accurate and time-efficient virtual dissection and analysis of large-scale tractography datasets* (p. 2025.05.27.656234). bioRxiv. https://doi.org/10.1101/2025.05.27.656234

Dell'Acqua, F., Lacerda, L. M., Catani, M., & Simmons, A. (2013). *Anisotropic Power Maps: A diffusion contrast to reveal low anisotropy tissues from HARDI data*. https://api.semanticscholar.org/CorpusID:46663779

Dell'acqua, F., Scifo, P., Rizzo, G., Catani, M., Simmons, A., Scotti, G., & Fazio, F. (2010). A modified damped Richardson-Lucy algorithm to reduce isotropic background effects in spherical deconvolution. *NeuroImage*, *49*(2), 1446–1458. https://doi.org/10.1016/j.neuroimage.2009.09.033

Dell'Acqua, F., Simmons, A., Williams, S. C. R., & Catani, M. (2013). Can spherical deconvolution provide more information than fiber orientations? Hindrance modulated orientational anisotropy, a true-tract specific index to characterize white matter diffusion. *Human Brain Mapping*, *34*(10), 2464–2483. https://doi.org/10.1002/hbm.22080

Dronkers, N. F. (1996). A new brain region for coordinating speech articulation. *Nature*, *384*(6605), 159–161. https://doi.org/10.1038/384159a0

Dronkers, N. F. (2000). The Pursuit of Brain–Language Relationships. *Brain and Language*, *71*(1), 59–61. https://doi.org/10.1006/brln.1999.2212

Eickhoff, S. B., Heim, S., Zilles, K., & Amunts, K. (2009). A systems perspective on the effective connectivity of overt speech production. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, *367*(1896), 2399–2421. https://doi.org/10.1098/rsta.2008.0287

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., & Gorgolewski, K. J. (2019). fMRIPrep: A robust preprocessing pipeline for functional MRI. *Nature Methods*, *16*(1), 111–116. https://doi.org/10.1038/s41592-018-0235-4

Faillenot, I., Heckemann, R. A., Frot, M., & Hammers, A. (2017). Macroanatomy and 3D probabilistic atlas of the human insula. *NeuroImage*, *150*, 88–98. https://doi.org/10.1016/j.neuroimage.2017.01.073

Fairbanks, G. (1955). Selective vocal effects of delayed auditory feedback. *The Journal of Speech and Hearing Disorders*, *20*(4), 333–346. https://doi.org/10.1044/jshd.2004.333

Fonov, V., Evans, A. C., Botteron, K., Almli, C. R., McKinstry, R. C., Collins, D. L., & Brain Development Cooperative Group. (2011). Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage*, *54*(1), 313–327. https://doi.org/10.1016/j.neuroimage.2010.07.033

Forkel, S. J., Bortolami, C., Dulyan, L., Barrett, R. L. C., & Beyh, A. (2025). Chapter 21—Dissecting white matter pathways: A neuroanatomical approach. In F. Dell'Acqua, M.

Descoteaux, & A. Leemans (Eds.), *Handbook of Diffusion MR Tractography* (pp. 397–421). Academic Press. https://doi.org/10.1016/B978-0-12-818894-1.00012-4

Forkel, S. J., Labache, L., Nachev, P., Thiebaut de Schotten, M., & Hesling, I. (2022). Stroke disconnectome decodes reading networks. *Brain Structure & Function*, *227*(9), 2897–2908. https://doi.org/10.1007/s00429-022-02575-x

Forkel, S. J., Rogalski, E., Drossinos Sancho, N., D'Anna, L., Luque Laguna, P., Sridhar, J., Dell'Acqua, F., Weintraub, S., Thompson, C., Mesulam, M.-M., & Catani, M. (2020). Anatomical evidence of an indirect pathway for word repetition. *Neurology*, *94*(6), e594–e606. https://doi.org/10.1212/WNL.0000000000008746

Forkel, S. J., Thiebaut de Schotten, M., Dell'Acqua, F., Kalra, L., Murphy, D. G. M., Williams, S. C. R., & Catani, M. (2014). Anatomical predictors of aphasia recovery: A tractography study of bilateral perisylvian language networks. *Brain*, *137*(7), 2027–2039. https://doi.org/10.1093/brain/awu113

Franken, M. K., Eisner, F., Acheson, D. J., McQueen, J. M., Hagoort, P., & Schoffelen, J.-M. (2018). Self-monitoring in the cerebral cortex: Neural responses to small pitch shifts in auditory feedback during speech production. *NeuroImage*, *179*, 326–336. https://doi.org/10.1016/j.neuroimage.2018.06.061

Friederici, A. D. (2009). Pathways to language: Fiber tracts in the human brain. *Trends in Cognitive Sciences*, *13*(4), 175–181. https://doi.org/10.1016/j.tics.2009.01.001

Glasser, M. F., & Rilling, J. K. (2008). DTI tractography of the human brain's language pathways. *Cerebral Cortex (New York, N.Y.: 1991)*, *18*(11), 2471–2482. https://doi.org/10.1093/cercor/bhn011

Goldberg, T. E., Gold, J. M., Coppola, R., & Weinberger, D. R. (1997). Unnatural practices, unspeakable actions: A study of delayed auditory feedback in schizophrenia. *The*

American Journal of Psychiatry, *154*(6), 858–860. https://doi.org/10.1176/ajp.154.6.858

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

Hertrich, I., Dietrich, S., & Ackermann, H. (2016). The role of the supplementary motor area for speech and language processing. *Neuroscience & Biobehavioral Reviews*, *68*, 602–610. https://doi.org/10.1016/j.neubiorev.2016.06.030

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews. Neuroscience*, *8*(5), 393–402. https://doi.org/10.1038/nrn2113

Houde, J. F., & Nagarajan, S. S. (2011). Speech production as state feedback control. *Frontiers in Human Neuroscience*, *5*, 82. https://doi.org/10.3389/fnhum.2011.00082

Howell, P., & Archer, A. (1984). Susceptibility to the effects of delayed auditory feedback. *Perception & Psychophysics*, *36*(3), 296–302. https://doi.org/10.3758/bf03206371

Irfanoglu, M. O., Nayak, A., Jenkins, J., & Pierpaoli, J. (n.d.). *(ISMRM 2017) TORTOISE v3: Improvements and New Features of the NIH Diffusion MRI Processing Pipeline*. Retrieved September 17, 2025, from https://archive.ismrm.org/2017/3540.html

Ivanova, M. V., Zhong, A., Turken, A., Baldo, J. V., & Dronkers, N. F. (2021). Functional Contributions of the Arcuate Fasciculus to Language Processing. *Frontiers in Human Neuroscience*, *15*, 672665. https://doi.org/10.3389/fnhum.2021.672665

Janacsek, K., Shattuck, K. F., Tagarelli, K. M., Lum, J. A. G., Turkeltaub, P. E., & Ullman, M. T. (2020). Sequence learning in the human brain: A functional neuroanatomical meta-analysis of serial reaction time studies. *NeuroImage*, *207*, 116387. https://doi.org/10.1016/j.neuroimage.2019.116387

Kalinowski, J., & Stuart, A. (1996). Stuttering amelioration at various auditory feedback delays and speech rates. *European Journal of Disorders of Communication: The Journal of the College of Speech and Language Therapists, London*, *31*(3), 259–269. https://doi.org/10.3109/13682829609033157

Kellner, E., Dhital, B., Kiselev, V. G., & Reisert, M. (2016). Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magnetic Resonance in Medicine*, *76*(5), 1574–1581. https://doi.org/10.1002/mrm.26054

Kort, N. S., Nagarajan, S. S., & Houde, J. F. (2014). A bilateral cortical network responds to pitch perturbations in speech feedback. *NeuroImage*, *86*, 525–535. https://doi.org/10.1016/j.neuroimage.2013.09.042

Kurth, F., Zilles, K., Fox, P. T., Laird, A. R., & Eickhoff, S. B. (2010). A link between the systems: Functional differentiation and integration within the human insula revealed by meta-analysis. *Brain Structure and Function*, *214*(5), 519–534. https://doi.org/10.1007/s00429-010-0255-z

Lee, B. S. (1950). Effects of delayed speech feedback. *Journal of the Acoustical Society of America*, *22*, 824–826. https://doi.org/10.1121/1.1906696

Lin, I.-F., Mochida, T., Asada, K., Ayaya, S., Kumagaya, S.-I., & Kato, M. (2015). Atypical delayed auditory feedback effect and Lombard effect on speech production in high-functioning adults with autism spectrum disorder. *Frontiers in Human Neuroscience*, *9*, 510. https://doi.org/10.3389/fnhum.2015.00510

Marien, P., Pickut, B. A., Engelborghs, S., Martin, J.-J., & De Deyn, P. P. (2001). Phonological agraphia following a focal anterior insulo-opercular infarction. *Neuropsychologia*, *39*(8), 845–855. https://doi.org/10.1016/S0028-3932(01)00006-9

Markiewicz, C. J., Esteban, O., Goncalves, M., Provins, C., Salo, T., Kent, J. D., DuPre, E., Ciric, R., Pinsard, B., Blair, R. W., Poldrack, R. A., & Gorgolewski, K. J. (2025).

fMRIPrep: A robust preprocessing pipeline for functional MRI [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.16657003

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

Munhall, K. G., MacDonald, E. N., Byrne, S. K., & Johnsrude, I. (2009). Talkers alter vowel production in response to real-time formant perturbation even when instructed not to compensate. *The Journal of the Acoustical Society of America*, *125*(1), 384–390. https://doi.org/10.1121/1.3035829

Niziolek, C. A., & Guenther, F. H. (2013). Vowel category boundaries enhance cortical and behavioral responses to speech feedback alterations. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(29), 12090–12098. https://doi.org/10.1523/JNEUROSCI.1008-13.2013

Ohlerth, A.-K., Lavrador, J. P., Vergani, F., & Forkel, S. J. (2025). Combining nTMS and Tractography for Language Mapping: An Integrated Paradigm for Neurosurgical Planning. In S. M. Krieg & T. Picht (Eds.), *Navigated Transcranial Magnetic Stimulation in Neurosurgery* (pp. 185–213). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-97155-6_10

Ozker, M., Doyle, W., Devinsky, O., & Flinker, A. (2022). A cortical network processes auditory error signals during human speech production to maintain fluency. *PLoS Biology*, *20*(2), e3001493. https://doi.org/10.1371/journal.pbio.3001493

Ozker, M., & Hagoort, P. (2025). Susceptibility to auditory feedback manipulations and individual variability. *PloS One*, *20*(5), e0323201. https://doi.org/10.1371/journal.pone.0323201

Ozker, M., Yu, L., Dugan, P., Doyle, W., Friedman, D., Devinsky, O., & Flinker, A. (2024). Speech-induced suppression and vocal feedback sensitivity in human cortex. *eLife*, *13*, RP94198. https://doi.org/10.7554/eLife.94198

Quabs, J., Bittner, N., & Caspers, S. (2025). Structural Connectivity Differences Reflect Microstructural Heterogeneity of the Human Insular Cortex. *Human Brain Mapping*, *46*(8), e70231. https://doi.org/10.1002/hbm.70231

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences*, *98*(2), 676–682. https://doi.org/10.1073/pnas.98.2.676

Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nature Neuroscience*, *12*(6), 718–724. https://doi.org/10.1038/nn.2331

Rojkova, K., Volle, E., Urbanski, M., Humbert, F., Dell'Acqua, F., & Thiebaut de Schotten, M. (2016). Atlasing the frontal lobe connections and their variability due to age and education: A spherical deconvolution tractography study. *Brain Structure & Function*, *221*(3), 1751–1766. https://doi.org/10.1007/s00429-015-1001-3

Seidler, R. D., Purushotham, A., Kim, S.-G., Ugurbil, K., Willingham, D., & Ashe, J. (2005). Neural correlates of encoding and expression in implicit sequence learning. *Experimental Brain Research*, *165*(1), 114–124. https://doi.org/10.1007/s00221-005-2284-z

Stuart, A., Kalinowski, J., Rastatter, M. P., & Lynch, K. (2002). Effect of delayed auditory feedback on normal speakers at two speech rates. *The Journal of the Acoustical Society of America*, *111*(5 Pt 1), 2237–2241. https://doi.org/10.1121/1.1466868

Takaso, H., Eisner, F., Wise, R. J., & 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: JSLHR*, *53*(2), 226–236. https://doi.org/10.1044/1092-4388(2009/09-0009)

Thiebaut de Schotten, M., Dell'Acqua, F., Valabregue, R., & Catani, M. (2012). Monkey to human comparative anatomy of the frontal lobe association tracts. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *48*(1), 82–96. https://doi.org/10.1016/j.cortex.2011.10.001

Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes*, *26*(7), 952–981. https://doi.org/10.1080/01690960903498424

Toyomura, A., Koyama, S., Miyamaoto, T., Terao, A., Omori, T., Murohashi, H., & Kuriki, S. (2007). Neural correlates of auditory feedback control in human. *Neuroscience*, *146*(2), 499–503. https://doi.org/10.1016/j.neuroscience.2007.02.023

Tremblay, S., Shiller, D. M., & Ostry, D. J. (2003). Somatosensory basis of speech production. *Nature*, *423*(6942), 866–869. https://doi.org/10.1038/nature01710

Veraart, J., Fieremans, E., & Novikov, D. S. (2016). Improvements and new features of the NIH diffusion MRI processing pipeline. *Magnetic Resonance in Medicine*, *76*(5), 1582–1593. https://doi.org/10.1002/mrm.26059

Weerathunge, H. R., Voon, T., Tardif, M., Cilento, D., & Stepp, C. E. (2022). Auditory and somatosensory feedback mechanisms of laryngeal and articulatory speech motor control. *Experimental Brain Research*, *240*(7), 2155–2173. https://doi.org/10.1007/s00221-022-06395-7