Cortical tracking of speech is reduced in adults who stutter when listening for speaking

Simone Gastaldon [1,2] *°, Pierpaolo Busan [3] °, Nicola Molinaro [4,5], Mikel Lizarazu [4]

1) Department of Developmental and Social Psychology (DPSS), University of Padova, Padova, Italy
2) Padova Neuroscience Center (PNC), University of Padova, Padova, Italy
3) Department of Medical, Surgical and Health Sciences, University of Trieste, Trieste, Italy
4) Basque Center on Cognition, Brain and Language (BCBL), Donostia-San Sebastián, Spain
5) Ikerbasque, Basque Foundation for Science, Bilbao, Spain

° Equal contribution

*** Corresponding author:**

Simone Gastaldon
Dipartimento di Psicologia dello Sviluppo e della Socializzazione (DPSS)
Università di Padova
Via Venezia 8, 35131, Padova (PD), Italy
E-mail: simone.gastaldon@unipd.it

ORCIDs:

Simone Gastaldon: 0000-0002-0818-2131
Pierpaolo Busan: 0000-0002-3847-5883
Nicola Molinaro: 0000-0002-7549-6042
Mikel Lizarazu: 0000-0001-5660-3087

This is a non-peer-reviewed manuscript. The final published version may partially differ in some respects. Please, check the bioRxiv page for updates on the published version.

Abstract

This study explores cortical tracking of speech (CTS) in adults who stutter (AWS) compared to typically fluent adults (TFA) while listening to sentences. We manipulated the upcoming involvement of the speech-motor system during listening: participants either had to simply listen to the sentences (listening only) or complete unfinished sentences by naming a picture (listening-for-speaking). AWS, known for atypical neural structure and functionaing in the speech-motor network, exhibited reduced CTS in the theta band in temporal sensors during the listening-for-speaking task, reflected at the source level in the left temporo-parietal junction and the right pre-motor and supplementary motor regions. Additionally, connectivity analyses reveal that TFA had stronger inter- and intra-hemispheric information transfer in the theta range than AWS in both tasks, involving frontal, temporo-parietal, (pre-)motor, and superior temporal regions, with different patterns according to the task. Notably, increased connectivity from the right superior temporal cortex to the left sensorimotor cortex correlated with faster naming times in the listening-for-speaking task. These findings suggest that atypical speech-motor functioning in stuttering impact also speech perception, especially in situations requiring articulatory alertness, and highlight the involvement of frontal and (pre-) motor regions in normal conditions in CTS.

Significance Statement

This study shows for the first time that individuals with a speech-motor impairment, namely persistent developmental stuttering, also show impaired cortical tracking of speech, especially when upcoming speech production is required. The effects emerge in the theta range, corresponding to the syllabic rhythm, suggested to be an optimal interface between the human biomechanic constraints for producing sounds and the human brain's perceptual capabilities for speech. Our study highlights the relevance of speech-motor regions in cortical tracking of speech and suggests that spoken language perception in individuals with speech-motor deficits is an important ground for research, especially in real-life conversational settings where smooth transitioning between listening and speaking is required.

Introduction

Developmental Stuttering (DS, also known as Childhood-Onset Fluency Disorder; American Psychiatric Association, 2013) is a neurodevelopmental disorder of the normal flow of speech, characterized by symptoms such as blocks, prolongations, and repetitions. People who stutter know what they want to say, but they may be unable to speak in a fluent manner. Importantly, DS may persist also in adulthood, impairing quality of life of affected individuals (Craig et al., 2009; Nang et al., 2018).

DS has a likely multifactorial origin, comprising genetic factors (Barnes et al., 2016; Benito-Aragón et al., 2020; Chow et al., 2020; Frigerio-Domingues & Drayna, 2017; Kang et al., 2010; Kang & Drayna, 2012; Kraft & Yairi, 2011) that may facilitate the appearance of a series of abnormal neural patterns (Alm, 2021a; Craig-McQuaide et al., 2014; A. C. Etchell et al., 2018). In this context, affected neural functions may result in impaired capacities related to the sensorimotor planning and execution of speech (Alm, 2021b; Chang et al., 2019), also affecting sensorimotor brain rhythms (A. C. Etchell et al., 2016; Ghaderi et al., 2018; Jenson et al., 2018, 2020; Joos et al., 2014; Saltuklaroglu et al., 2017). As a matter of fact, DS seems to be characterized by the presence of a deficit in internal timing and motor coordination (Alm, 2004), involving wide neural systems and comprising regions such as the basal ganglia, the supplementary motor area, the inferior frontal cortex, and temporal regions (Busan et al., 2019; Busan, 2020; Craig-McQuaide et al., 2014; A. Etchell et al., 2018; Watkins et al., 2008). Importantly, neural inefficiency that perturbs speech-motor execution seems to also affect aspects of speech comprehension, specifically leading to weaker or less efficient predictive processing (Gastaldon et al., 2023). Compatibly, audio-motor interactions seem to be bi-directionally impaired in DS (Chang et al., 2016; Daliri & Max, 2015, 2018; Halag-Milo et al., 2016; Kim et al., 2020). From a behavioral point of view, malfunctioning of predictive timing during auditory-motor coupling in people who stutter can lead to some differences when compared to fluent speakers (e.g., Falk et al., 2015; however, works reporting no overt differences in auditory-motor integration are also available, e.g., Assaneo et al., 2022).

Audio-motor coupling is fundamental for driving effective speech production, but evidence suggests that this is relevant also for speech perception (Poeppel & Assaneo, 2020). Cortical tracking of speech (CTS, sometimes also referred to “speech-brain entrainment”; Assaneo & Poeppel, 2018; Molinaro & Lizarazu, 2018; Poeppel & Assaneo, 2020; Poeppel & Teng, 2020) is the ability to track quasi-rhythmic information in the flow of speech by aligning the phase of internal brain rhythms at low frequency to the envelope (acoustic energy) of the perceived sentences, and is taken to be a valuable index reflecting the efficiency of neural processing of speech. CTS mainly results in auditory neural tracking in the delta/theta frequency bands, with a peak usually between 2 and 8 Hz (Poeppel & Assaneo, 2020). Interestingly, a body of evidence supports the view that frontal, motor and premotor regions modulate, in a top-down direction, neural tracking at low frequencies in the auditory cortex (Keitel et al., 2018; Park et al., 2015). Evidence also suggests that there is a preferred frequency range at which motor and auditory cortices are coupled, i.e., in the theta range around 4.5 Hz (Assaneo & Poeppel, 2018). In this scenario, the motor system is supposedly exploited to reduce noise and uncertainty, by generating temporal predictions via efferent motor signals, causing phase-resetting in auditory cortices and optimizing sensory perception (Rimmele et al., 2018). Recent behavioral evidence supports this account: individual speech production rates and auditory-motor coupling explain performance in a speech comprehension task (higher speech rates and auditory-motor synchronization are associated with higher comprehension scores; Lubinus et al., 2023).

Importantly, altered brain processes related to CTS have been proposed as a risk factor for the appearance of developmental speech and/or language disorders (usually characterized by abnormalities in cortico-basal-thalamo-cortical circuitries involved in the processing of sensory cues - such as beats in music and/or linguistic meter in speech - thus interacting with auditory stimuli and playing a role in processing and predicting events in a sequence; Ladányi et al., 2020). For example, CTS is abnormally modulated in people with dyslexia (Lizarazu et al., 2015; Molinaro et al., 2016). This may be the case also for people who stutter, especially if auditory-motor coupling is a contributory factor to both DS and CTS. Crucially, no evidence is currently available for brain processes related to CTS in stuttering. Thus, a better understanding of these phenomena in DS should be useful for 1) improving our comprehension of neuro-pathological mechanisms related to stuttering (both in the contexts of speech production and perception), and 2) improving our comprehension of the neural mechanisms involved in typical speech perception and production (and in their possible mutual interactions).

To fill this gap of knowledge, in the present study we investigated whether adults who stutter (AWS; stuttering onset during childhood and persisting into adulthood) show altered tracking of the speech signal when compared to typically fluent adults (TFA; no diagnosis of speech disorders). Specifically, we measured speech-brain coherence on electroencephalographic (EEG) data, both at the sensor and the neural source level, during sentence listening in conditions that either overtly recruited the articulatory system (completing the sentence by naming a picture; listening-for-speaking) or not (listen the entire sentence; listening only), in order to better understand if the upcoming involvement of the speech-motor network may have modulatory effects on CTS (see Figure 1 and Materials and Methods). While not directly simulating everyday two-people conversations, the listening-for-speaking task still implies an alertness of the speech-motor system, in addition to higher level processes such as anticipation and planning (Corps et al., 2018), similar to what is required in conversational and turn-taking settings.

In addition, partial directed coherence (PDC) was analyzed to quantify directional neural interactions between brain regions implicated in sensorimotor speech processing. This analysis allows us to investigate frequency-specific directional communication between cortical regions during envelope tracking, possibly highlighting differences in speech-relevant pathways. By taking as reference Hickok and Poeppel's dual route model of speech processing (Hickok & Poeppel, 2004, 2007), we may hypothesize that neural connectivity is reduced in stuttering especially in the dorsal stream, suggested to be responsible for motor-auditory transformations (Hickok et al., 2011). This hypothesis is also supported by evidence showing that white matter tracts considered to be part of the dorsal stream are usually altered in people who stutter (Kronfeld-Duenias et al., 2016; Neef et al., 2018a, 2022; Sommer et al., 2002; Watkins et al., 2008). Thus, given the picture outlined above, we can hypothesize that:

1. CTS may be reduced in AWS relative to TFA, regardless of the listening condition (listening only vs listening-for-speaking). Alternatively, differences may be detected only when listening is coupled with the upcoming necessity to overtly activate the speech-motor system, which is consequently kept in an "alert mode" in order to appropriately initiate speech.
2. In AWS, reduced CTS may be found in speech-motor and premotor regions, in addition to auditory regions.
3. In AWS, regions that are considered to be part of the dorsal stream (inferior frontal cortex, premotor and supplementary motor regions, sensorimotor and temporo-parietal regions) may be communicating less efficiently with auditory regions during speech tracking, thus displaying reduced connectivity.

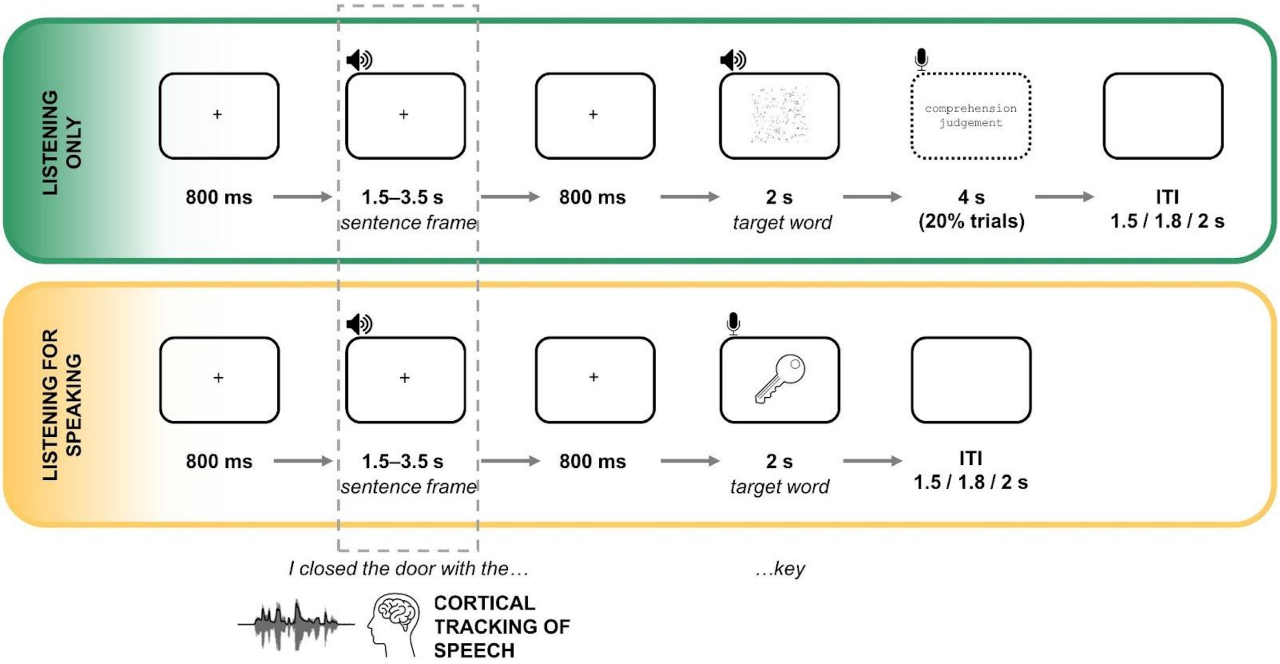


Figure 1. Experimental design. Participants listened to sentence frames and target final words in the listening-only task, while they had to complete the sentence frame by naming a picture (representing the target word) in the listening-for-speaking task. Indexes related to cortical tracking of speech and cortical connectivity were measured during the auditory presentation of the sentence frames. Response times in the listening-for-speaking task were collected (calculated at picture onset).

Materials and Methods

Participants

We analyzed CTS in a dataset collected for a previous study on spoken sentence processing in adults who stutter which focused on different time-windows, different analyses and different aims (Gastaldon et al., 2023). The original study included 14 right-handed male adults who stutter (AWS) and 14 right-handed typically fluent male adults (TFA). The participants were matched for age and handedness. The original study was approved by the Ethical Committee for Psychological Research of the University of Padova and conducted in accordance with the Declaration of Helsinki. We refer the reader to the original study for further details on AWS recruitment and assessment. Out of the 28 participants of the original study, 24 were analyzed in the present work: 12 AWS (mean age = 34.44, SD = 9.37) and 12 TFA (mean age = 33.42, SD = 8.94). The remaining participants were also matched for handedness, as assessed by means of the Edinburgh Handedness Inventory (Oldfield, 1971): AWS: mean = 83.75, SD = 20.57; TFA: mean = 85.00, SD = 23.06. Four participants were excluded due to excessive noisy EEG data for the present type of analysis during sentence frame presentation. From the original study we also retrieved data about the Stuttering Severity Index (SSI-4; Riley, 2009) of each participant, in order to perform correlations with CTS and connectivity data (statistical threshold for explorative correlations: $p \leq 0.01$, “two-tailed”).

Stimuli and procedure

The stimuli were the same used for (Gastaldon et al., 2020, 2023); see the OSF repository for additional information on stimuli characteristics: https://osf.io/tcbsh/). They consisted of 256 sentence frames (sentences without the last word, ranging in duration from 1.55 to 3.54 s; mean duration = 2.39, SD = 0.4), 128 target words and 128 b/w line pictures (124x124 pixels). Each word and picture were coupled with 2 of the sentence frames. Spoken stimuli were uttered by an Italian native female speaker, recorded and digitized at 44.1 kHz using Audacity[1]. Audio files (*.wav) were also segmented using Audacity. During the task, as sentence completion, participants either heard a word (listening-only task) or had to produce it by naming a picture (listening-for-speaking task), in two distinct blocks. For both listening conditions, sentences could either be highly or low constraining towards a specific word completing the sentence. In this study, in order to allow for a better estimation of CTS and increase statistical power, we have not divided the unfinished sentence into high vs low constraining as in the original study (Gastaldon et al., 2023). Instead, we focused on the manipulation of the task, which implied two different listening conditions: listen with the aim of comprehending and answer to a true/false judgment question that appeared on the screen in the 20% of the trials (listening only), and listen in order to be ready to complete the sentence as quickly as possible by naming a picture (listening-for-speaking). Participants sit in a dimly lit room in front of a computer screen. The experimental material was delivered through E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA). Unfinished sentences were presented through built-in speakers. Responses (true/false answers and naming) were collected via a microphone set in front of the participant. In the listening-for-speaking task, audio recording started at the onset of the picture to be named and lasted for 2 seconds.

---

[1] Audacity(R) software is copyright (c) 1999-2014 Audacity Team. The name Audacity(R) is a registered trademark of Dominic Mazzoni.

The experimental paradigm is exemplified in Figure 1. For further details on the experimental design, we refer the reader to the original study (Gastaldon et al., 2023).

EEG data acquisition and preprocessing

During the task, the electroencephalogram (EEG) was recorded using a BrainAmp amplifier and BrainVision Recorder software (BrainProducts, Germany). EEG was recorded using 64 electrodes that were positioned according to the international 10-10 system (Nuwer et al., 1998). Scalp-electrode impedance was kept below 10Ω. The recording was referenced to the left earlobe. Electrode AFz served as the ground. Two electrodes at the outer canthi of both eyes recorded horizontal eye movements and one electrode below the left eye recorded vertical eye movements. EEG was sampled at 1000 Hz and band-pass filtered online from 0.1 to 1000 Hz.

The preprocessing pipeline for the present works was the following. Heartbeat and EOG artifacts were identified using independent component analysis (ICA) and subtracted from the recordings in a linear manner. The ICA decomposition was carried out using the Infomax algorithm implemented in the Fieldtrip toolbox (Oostenveld et al., 2011). Across participants, the number of heartbeat and ocular components that were removed varied from 1 to 4 and 1 to 3 components, respectively. Furthermore, trials were visually inspected to discard any remaining artifacts. Bad channels were substituted with interpolated values computed as the average of the neighboring electrodes obtained through the triangulation method implemented in Fieldtrip. A minimum of 75% artifact-free trials per participant was required for inclusion in subsequent analyses. Two participants from each group were below this threshold and have been discarded from further analyses, thus resulting in a final sample of 24 participants (12 AWS and 12 TFA). In the case of TFA, an average of 4.89% (SD = 3.1) trials and 5.23% (SD = 3.8) trials were excluded for comprehension and production tasks, respectively. Similarly, for AWS, an average of 5.4% (SD = 3.59) trials and 8.13% (SD = 3.86) trials were excluded for comprehension and production tasks, respectively. Importantly, no significant group or task differences were observed in the number of excluded trials (all Ts < 1.6, all ps > 0.11, two-tailed t-test). EEG data and Matlab scripts for the analyses described in the following paragraphs are available on a dedicated OSF repository: https://osf.io/7gpyb/.

Cortical tracking of speech (CTS) analysis

Sensor level analysis

We used coherence to measure the coupling between the speech envelope and the cortical activity, here referred to as the cortical tracking of speech (CTS). We obtained the envelope of the speech signal from the Hilbert transformed broadband stimulus waveform. For each EEG sensor, coherence between the speech envelope and the artifact free epochs was calculated in the 1 – 15 Hz frequency band with 1 Hz frequency resolution (Molinaro et al., 2016; Molinaro & Lizarazu, 2018). This procedure was followed for each participant and task/listening condition.

According to the previous research in speech processing we expect to find strong CTS in the low-frequency (< 10 Hz) spectrum and in temporal sensors (Lizarazu et al., 2021; Molinaro et al., 2016; Molinaro & Lizarazu, 2018). Therefore, we selected a set of 12 channels, evenly distributed to cover the temporal lobes of the brain – precisely, 6 channels allocated to the left hemisphere (C3, C5, CP3, CP5, FC3, FC5) and additional 6 to the right hemisphere (C4, C6, CP4, CP6, FC4, FC6). To estimate the coherence bias, the auditory envelopes were randomly shuffled across epochs for each participant, and coherence was recalculated in 100 permutations. The coherence data from the selected sensors of interest were separately averaged for each hemisphere and then transformed into z-scores using the mean and standard deviation derived from the 100 random EEG-audio pairings for those sensors. For each condition and frequency bin, z-score transformations were computed using the task-specific mean and standard deviation obtained from the random pairing dataset, and with an equal number of trials as the actual EEG-audio pairing dataset.

To assess group differences in each task, we conducted an ANOVA on the z-transformed coherence values, with hemisphere (left vs right) as the within-subject factor and group (TFA vs AWS) as the between-subject factor (considering effects of main factors and their interaction; post-hoc analyses conducted by using t-test; statistical threshold at $p \leq 0.05$, two-tailed).

Source level analysis

Coherence values were estimated also at the source level for each participant and condition at theta range (3-5 Hz), where significant results were observed at the scalp level. For that, we utilized a frequency-domain adaptive spatial filtering imaging of coherent sources algorithm (Gross et al., 2001), implemented in the Fieldtrip toolbox. To establish the spatial relationship between electrode positions of the participants (defined with a template electrode layout) and the cortical mesh, we employed a standard boundary element head model (BEM) extracted from the Montreal Neurological Institute (MNI) template. This BEM consists of three 3-D surfaces (skin, skull, brain) derived from the MNI dataset. The forward model was computed using an 8 mm grid encompassing the entire brain compartments of the BEM, representing various source positions. To perform source analysis, we constructed common space filters utilizing the leadfield of each grid point and the cross-spectral density matrix (CSD). The CSD matrices were computed within the theta (4 Hz with ±1 Hz frequency smoothing) band by applying the fast Fourier transform to 1-second data segments in sliding windows shifting in 0.5 seconds steps. As anticipated, the selection of the theta range was based on the observation of group effects at the sensor level occurring specifically at this frequency. Beamformer coefficients were computed considering the dominant source direction within all voxels and a regularization factor of 7% was applied. The coherence for each source location was estimated using the EEG data and the spatial filter in the theta band. To ensure comparability of source coherence values across subjects, we normalized individual coherence brain maps. For this reason, the coherence at each source was converted to a z-score value by subtracting the mean coherence across all sources and dividing by the standard deviation across all sources. Successively, for each group and condition, z-scored source coherence values were projected on the brain surface mesh image BrainMesh_ICBM152_smoothed from Surf Ice (Version 12.1; https://www.nitrc.org/projects/surfice/)

Finally, based on previous functional neural evidence on DS (Belyk et al., 2015, 2017; Brown et al., 2005b; Budde et al., 2014; Busan et al., 2019; Craig-McQuaide et al., 2014; A. C. Etchell et al., 2018; Ingham et al., 2012; Neef et al., 2015) we selected five regions of interest (ROIs) from the Automatic Anatomical Labeling (AAL; Tzourio-Mazoyer et al., 2002). As shown in Figure 2, within each cerebral hemisphere, we defined the subsequent brain regions: the inferior frontal gyrus (IFG, comprising the pars opercularis, pars triangularis, and pars orbitalis), the premotor and supplementary motor cortex (preM), the sensorimotor strip (SM, comprising the pre-central and post-central gyri), the inferior parietal lobule (IPL, comprising the supramarginal and angular gyri), and the superior temporal gyrus (STG).

For each task, we employed the Wilcoxon ranked sum non-parametric test to assess group differences on the mean of the z-scored coherence values within each ROI (statistical threshold at $p \leq 0.05$, two-tailed).

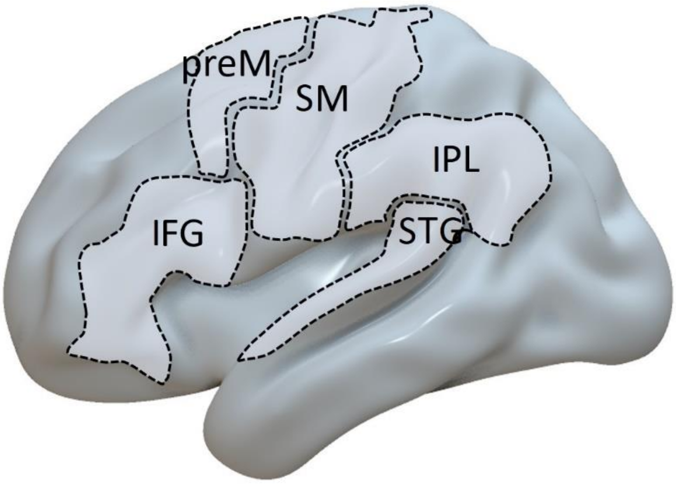


Figure 2. Regions of interest (ROIs) selected for statistical comparison between groups in the source and connectivity analyses. Five ROIs were selected in the left and right hemisphere: the inferior frontal gyrus (IFG, comprising the pars opercularis, triangularis, and orbitalis), the premotor and supplementary motor cortex (preM), the sensorimotor strip (SM), the inferior parietal lobule (IPL), and the superior temporal gyrus (STG).

Partial directed coherence (PDC) analysis

We employed partial directed coherence (PDC) to assess the causal connections between neural activity associated with speech processing within our designated ROIs (IFG, preM, SM, IPL and STG). After creating spatial filters, virtual time series in the source locations within the ROIs were reconstructed by applying the respective spatial filter to the EEG sensor data filtered in the theta (3 - 5 Hz) band. Because ROIs typically comprise many point sources, we employed principal component analysis (PCA) to identify the most representative time series within each ROI. To achieve this, we conducted a PCA on all time-series within each ROI and selected the first principal vector, which represented the distribution that explained most of the variance across all time-series that entered the PCA. For each participant and task, we computed PDC between the representative time series in each ROI for each participant and task. PDC is based on the Granger Causality principle (Granger, 1969; Seth et al., 2015) and on vector autoregressive (VAR) modeling of the data. The VAR model of order p for a variable x is given by:

$$x(t)=\sum_{r=1}^{p}a(r)s(t-r)+\varepsilon(t)$$

$$\begin{pmatrix} s_1(t) \\ \cdot \\ \cdot \\ \cdot \\ s_M(t) \end{pmatrix} = \sum_{r=1}^{p} a_r \begin{pmatrix} s_1(k-r) \\ \cdot \\ \cdot \\ \cdot \\ s_M(k-r) \end{pmatrix} + \begin{pmatrix} \varepsilon_1(t) \\ \cdot \\ \cdot \\ \cdot \\ \varepsilon_M(t) \end{pmatrix}$$

where s(t) = (s1(t), s2(t), ... , sM(t))Tare the stationary M-dimensional simultaneously measured time series in each ROI; ar are the M x M coefficient matrices of the model; and (t) is a multivariate Gaussian white noise process. In our case, M = 10 since we calculated the connectivity network formed by five different ROIs. The model order *p* was selected with the Schwartz Information Criterion. This criterion selects the model order that optimizes the goodness of fit of the model, while introducing a penalty depending on the complexity of the model. In the frequency domain the version of Granger-causality is given by:

$$A(f)=I-\sum_{r=1}^{p}a(r)e^{-i2\pi fr/p}$$

The first term of the difference refers to the identity matrix (M-dimensional) and the second one to the Fourier trans-form of the VAR coefficients. Then, the PDC from the ROI j to ROI i is given by:

$$PDC_{j\to i}(f)=\frac{|A_{ij}(f)|}{\sqrt{\sum_k |A_{kj}(f)|^2}}$$

The PDC provides a measure of the linear directional coupling strength of s_j on s_i at frequency *f* (theta). The PDC values vary between 0 (no directional coupling) and 1 (perfect directional coupling). PDC analysis was performed using the Frequency-Domain Multivariate Analysis toolbox (FDMa, Freiburg Center for Data Analysis and University of Freiburg, Germany), and the model order was computed using algorithms developed in the Multivariate Autoregressive Model Fitting (ARfit) software package (Schneider & Neumaier, 2001). For each task, we used the Wilcoxon ranked sum non-parametric test to assess group variations in directional coupling between ROIs (statistical threshold at p ≤ 0.05, two-tailed).

Naming accuracy and response times analysis (listening-for-speaking task)

For naming latencies (response times, RT), we took the data from (Gastaldon et al., 2023), also available here: https://osf.io/5jkur/. Here we summarize how latencies were derived in the original study, but we refer the reader to the original article for additional details. To estimate naming times, audio recordings (2 seconds *.wav files starting at picture onset) were fed to Chronset (Roux et al., 2017). Only correct responses were considered. Responses were coded as incorrect if they started with hesitation sounds, if corrections were made during the response, or if the participant could not produce enough of the target word in the 2-second recording (in order to be able to assess the correctness of the response).
Statistical analyses were performed in R. Accuracy was analyzed with a generalized linear mixed-effects model (GLMM) with binomial distribution family. Group, lexical frequency of the target word (retrieved through PhonItalia; Goslin et al., 2014) and repetition (the same target word was presented twice in the task, associated to two different sentence frames) were set as fixed effect, while participant and item as random intercepts. RTs were analyzed with a GLMM with gamma distribution family and identity link function. Group, lexical frequency of the target word and repetition were set as fixed effects, while participant and item as random intercepts. As explained above, we decided not to include constraint as a factor here since for the coherence analysis (the main focus of the present work) we did not differentiate between the two for methodological reasons. GLMM were fitted with the lme4 package (Bates et al., 2015) and contrasts set to sum coding. Finally, as for SSI-4, RTs were correlated with CTS and connectivity data (statistical threshold at $p \leq 0.01$, two-tailed).

Results

Naming (listening-for-speaking task)

Accuracy and response times (RTs) are shown in Figure 2, while model summaries are reported in Table 1. Participants of both groups had a very high accuracy in producing the correct word (AWS: mean = 0.96, SD = 0.2; TFA: mean = 0.99, SD = 0.1; see Figure 2A). However, the model revealed a main effect of repetition (higher accuracy when the picture appears for the second time) and a main effect of group, with AWS less accurate than TFA (see Table 1). Speculatively, however, rather than reflecting possible inefficiency in retrieving lexical items in AWS, this difference is likely due to the limited time available for recording the responses (2 seconds after picture onset): sometimes, AWS may have provided the correct response outside this window, hence the reduced accuracy (i.e., fewer trials coded as correct). In addition, as can also be seen from Figure 3A, this may be due to a subset of participants. For what concerns response times, AWS were slower than TFA (AWS: mean = 771.19 ms, SD = 267.77; TFA: mean = 650.53 ms, SD = 219.98; see Figure 3B). In fact, the model revealed a main effect of repetition and, importantly, a main effect of group (see Table 1).

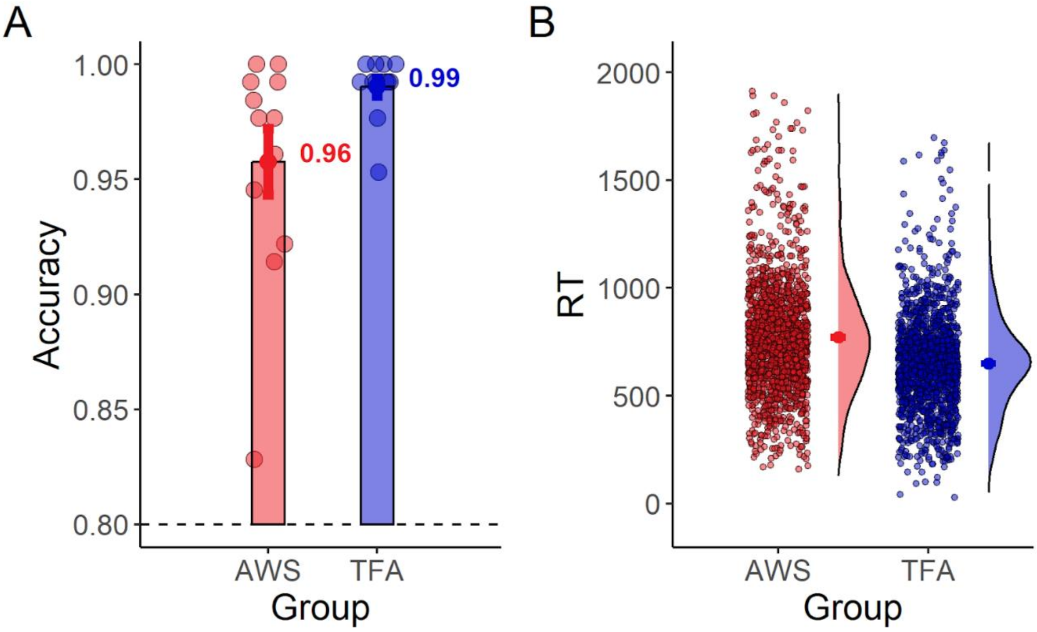


Figure 3. Behavioral results for the listening-for-speaking task. A) Subject-level accuracy scores (dots), and group-level accuracy means with error bars. B) Single-trial response times (dots), group-level means with error bars and density distributions.

Table 1. Model summaries for accuracy and RTs for naming in the listening-for-speaking task.

	ACCURACY			
Predictor	*Estimates*	*CI*	*Statistics*	*p-value*
(Intercept)	60.83	20.21 – 183.11	7.31	<0.001
lexical frequency	1.22	0.97 – 1.54	1.68	0.093
group	0.47	0.27 – 0.82	-2.64	0.008
repetition	0.70	0.55 – 0.90	-2.84	0.005
	RESPONSE TIMES			
Predictor	*Estimates*	*CI*	*Statistics*	*p-value*
(Intercept)	730.15	688.82 – 771.48	34.64	<0.001
lexical frequency	-6.05	-14.21 – 2.10	-1.45	0.146
group	60.43	34.44 – 86.41	4.56	<0.001
repetition	36.44	29.12 – 43.76	9.76	<0.001

Sensor-level CTS

Initially, we conducted an assessment of sensor-level cortical tracking of speech within the 1 - 15 Hz frequency range for each group (TFA and AWS) and task (listening-only and listening-for-speaking). Consistent with previous studies, we observed that during speech listening, CTS is highest in the theta (3 - 5 Hz) frequency band (Figure 4A) in bilateral temporal sensors (Figure 4B).

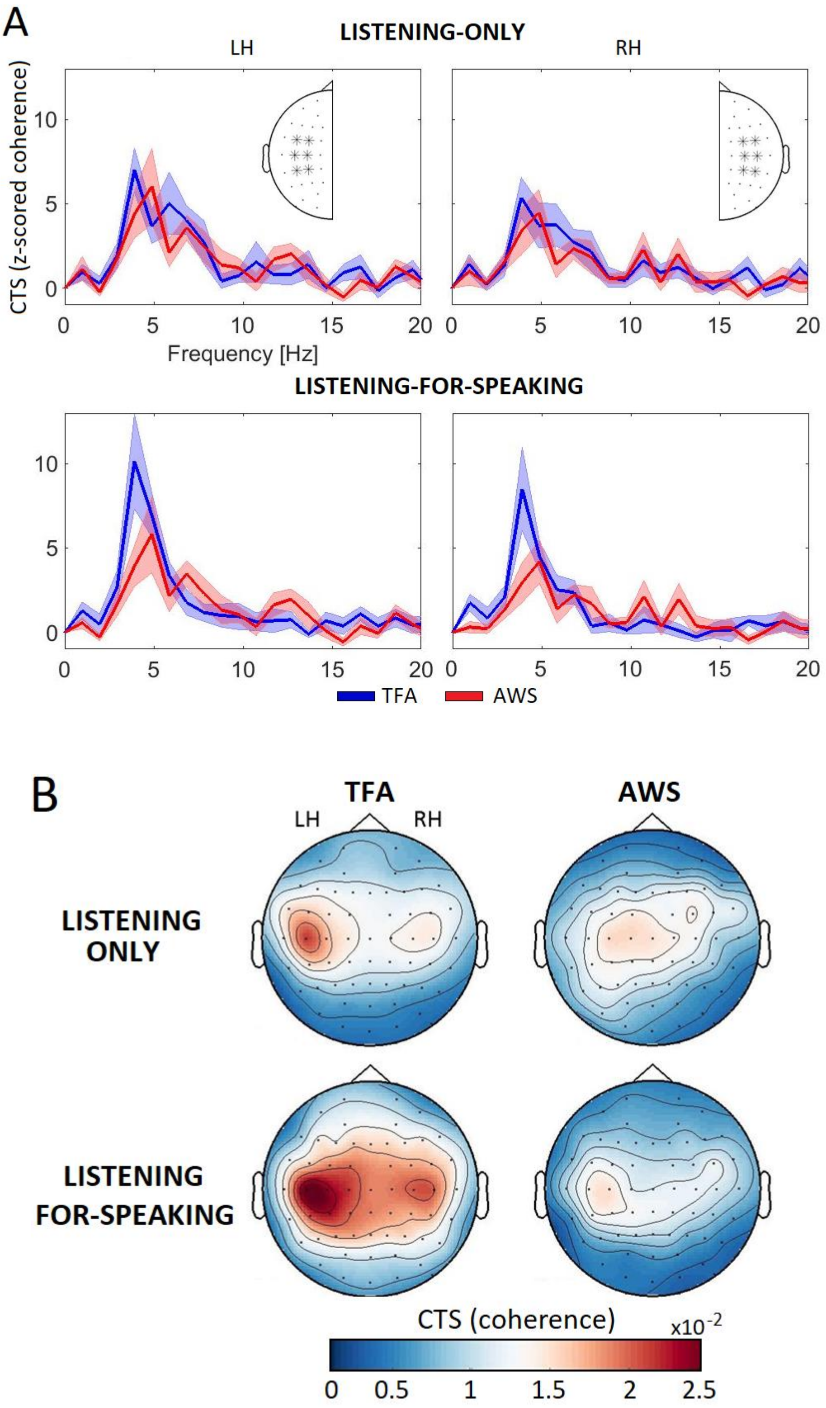


Figure 4. Spectra distribution and topographic map of the CTS at the sensor level. A) Corrected coherence values (coherence values converted into z-scores using the mean and standard deviation derived from the 100 random EEG-audio combinations) in the 1 – 15 Hz frequency range can be observed across representative sensors (C3, C5, CP3, CP5, FC3, FC5, C4, C6, CP4, CP6, FC4, FC6) of the left (LH) and right (RH) hemisphere. B) For each group (TFA: Typical Fluent Adults; AWS: Adults Who Stutter) and task (listening-only and listening-for-speaking), we represented the topographic maps of uncorrected coherence values in the theta (3 - 5 Hz) frequency band.

For each task, we performed an ANOVA on the mean CTS values (z-scored coherence) within the theta band and across the sensors of interest in both the left and right hemispheres. In the listening-only task, we did not observe any main effects or interactions in the CTS values (all $F(1,22) < 1.97$, all $p < 0.17$, $\eta 2 < 0.06$). However, we did observe a main effect of Group ($F(1,22) = 4.07$, $p = 0.05$, $\eta 2 = 0.15$) in the CTS values for the listening-for-speaking task. Post-hoc tests showed that CTS was significantly higher in TFA compared to AWS ($t = 2.02$, $p = 0.05$, Cohen's d = 0.80). No statistically significant correlations with RTs and SSI-4 were found.

Source-level CTS

When considering source analyses, we observed that for both the listening-only and the listening-for-speaking tasks, frontal, temporal, and parietal cortical regions showed strong CTS (z-scored coherence values) in the theta band (Figure 5). Subsequently, we calculated the mean of the CTS values in each of the ROIs described in the Materials and Methods section: the inferior frontal gyrus (IFG), the premotor/supplementary motor cortex (preM), the sensorimotor cortex (SM), the inferior parietal lobule (IPL), and the superior temporal gyrus (STG).

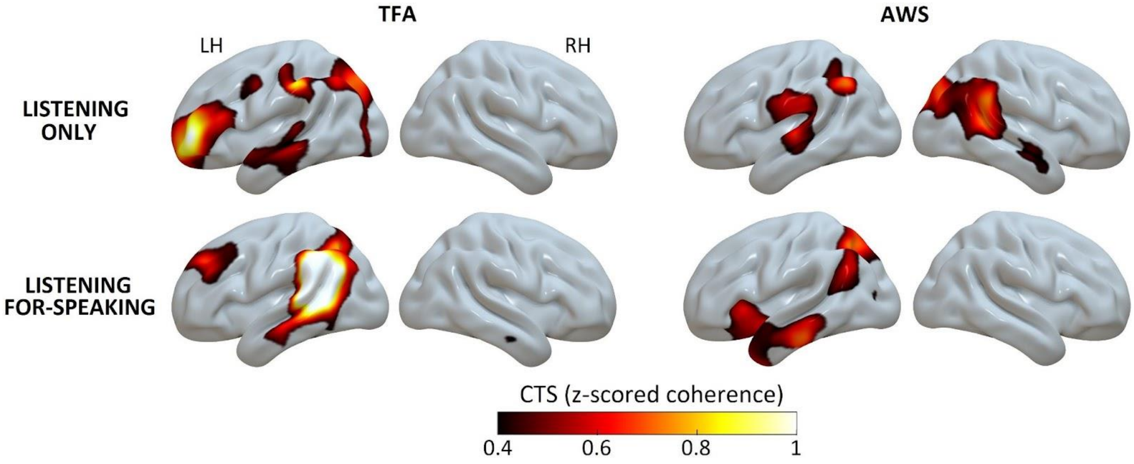


Figure 5. Source reconstruction of the CTS values in the theta range. For each group (TFA: Typical Fluent Adults; AWS: Adults Who Stutter) and listening condition (listening only and listening-for-speaking), we represented the source maps of CTS values (coherence values converted into z-scores using the mean and standard deviation derived from the CTS values in all the sources) in the theta (3 - 5 Hz) frequency band.

In line with the results observed at the sensor level, we found that the CTS values were significantly stronger for individuals with TFA compared to AWS in the left IPL ($M_{TFA} = 1.92$, $SD_{TFA} = 2.58$; $M_{AWS} = 0.22$, $SD_{AWS} = 0.34$; $p = 0.03$) and in the right preM regions ($M_{TFA} = 0.07$, $SD_{TFA} = 1.21$; $M_{AWS} = -0.74$, $SD_{AWS} = 0.36$; $p = 0.01$), only for the listening-for-speaking task (Figure 6). We did not find any group differences in the listening-only task (all $p > 0.09$) (Supplementary Figure 1). No statistically significant correlations with RTs and SSI-4 were found.

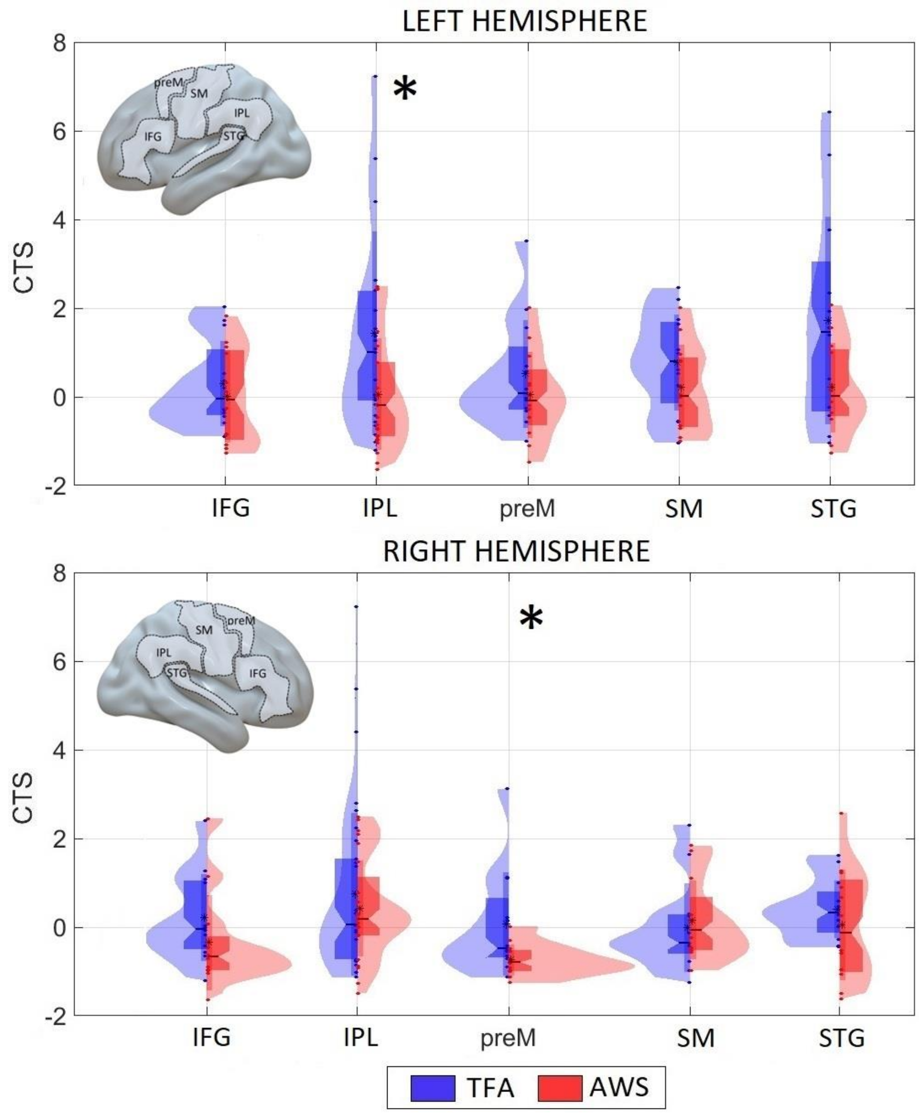


Figure 6. The mean CTS in the 3-5 Hz range (theta band) in each region of interest for each group. Boxplots are overlaid with individual data points. Each dot represents data from one of the participants (blue for TFA and red for AWS). Boxes cover the 25th to 75th percentile (inter-quartile range; IQR). The middle of the box represents the median. Whiskers extend from the 25th percentile and 75th percentile to cover all data points lying within 1.5 times the IQR (from the 25th and 75th percentile, respectively). Regions showing a significant group effect are marked with an asterisk.

Source-level connectivity

We used partial directed coherence (PDC) to assess causal functional connectivity during speech processing in the theta band (3 - 5 Hz) between different ROIs. For the listening-only task (Figure 7), we observed significantly higher connectivity in TFA compared to AWS from the left STG to the right IFG ($M_{TFA} = 0.12$, $SD_{TFA} = 0.05$; $M_{AWS} = 0.07$, $SD_{AWS} = 0.04$; $p = 0.01$), and from the right IFG to the left IPL ($M_{TFA} = 0.03$, $SD_{TFA} = 0.01$; $M_{AWS} = 0.02$, $SD_{AWS} < 0.01$; $p = 0.04$). For the listening-for-speaking task (Figure 5), we observed significantly higher connectivity in TFA compared to AWS from the right STG to the left IPL ($M_{TFA} = 0.05$, $SD_{TFA} = 0.06$; $M_{AWS} = 0.03$, $SD_{AWS} = 0.02$; $p = 0.05$) and from the right STG to the left SM regions ($M_{TFA} = 0.11$, $SD_{TFA} = 0.06$; $M_{AWS} = 0.06$, $SD_{AWS} = 0.04$; $p = 0.03$).

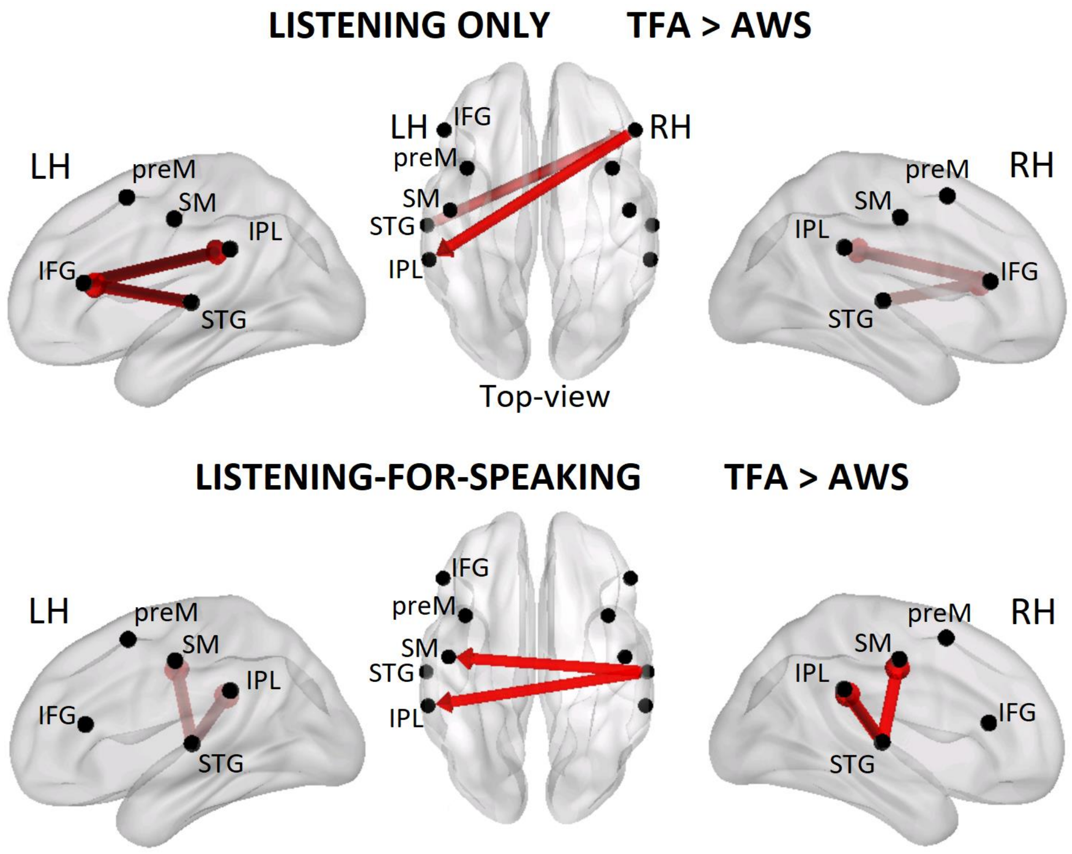


Figure 7. Causal functional connectivity analysis. For each task, we included the connections that exhibited statistically significantly higher PDC values for TFA compared to AWS. We included a seed for each of the regions of interest (IFG: inferior frontal gyrus; preM: premotor/supplementary motor; SM: sensorimotor; IPL: inferior parietal lobule; STG: superior temporal) in both the left (LH) and right (RH) hemisphere.

Interestingly, when considering both groups, a statistically significant negative correlation was found in the listening-for-speaking task between RTs and the connectivity from the right STG to the left SM cortex (r = -0.56, p = 0.0048): stronger directional connectivity between these regions is associated with faster response times (see Figure 8).

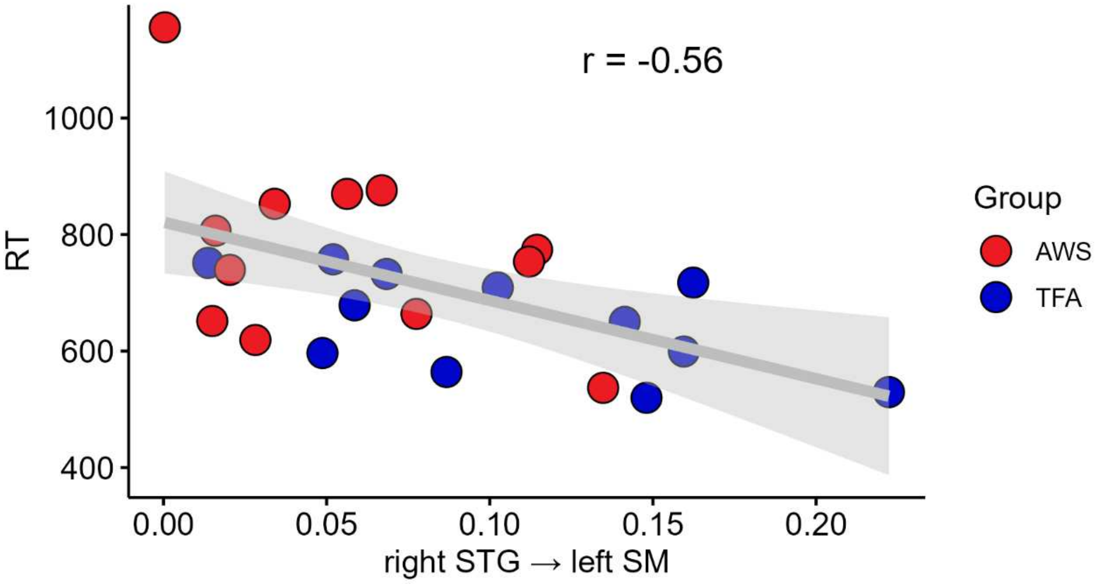


Figure 8. Scatterplot showing the correlation between response times (RTs) and connectivity from the right superior temporal gyrus to the left sensorimotor cortex.

Discussion

For the first time, in the present work we analyzed cortical tracking of speech (CTS) in a group of participants with dysfunctional speech-motor control due to a neurodevelopmental disorder, namely Developmental Stuttering (DS). We therefore compared adults who stutter (AWS) with typically fluent adults (TFA). To investigate the role of the alertness state of the speech-motor system on CTS, we analyzed two different listening situations: listening-only (no upcoming involvement of speech production) and listening-for-speaking (listen to an unfinished sentence and complete it by naming a picture; upcoming overt engagement of the speech-motor system). We observed reduced coherence in the theta range (3-5 Hz) in AWS relative to TFA in the listening-for-speaking task, both at the sensor (bilaterally around the temporal regions) and the neural source levels. More specifically, at source level, AWS resulted in lower CTS in the left inferior parietal/temporo-parietal cortex and in the right premotor and supplementary motor regions. Cortical connectivity measures in the theta range were differently modulated as well, with weaker connections in both listening conditions, primarily resulting in lower inter-hemispheric information exchange involving frontal, auditory/temporal and sensorimotor region. Notably, we also found slower (speech) response times in AWS in the listening-for-speaking task, and a significant negative correlation between one of the connectivity patterns and RTs in the listening-for-speaking task, reinforcing the arguments we lay out in the following.

Cortical tracking of speech at the syllabic rate is reduced in Developmental Stuttering when listening for speaking

As highlighted in the Introduction, CTS is a neural index reflecting speech-brain synchronization, resulting in the alignment of the phase of (internal) brain frequencies to acoustic features of the speech signal (Assaneo & Poeppel, 2018; Poeppel & Assaneo, 2020; Poeppel & Teng, 2020). Additionally, an intrinsic coupling with speech-motor production regions seems to support this process (Assaneo & Poeppel, 2018; Keitel et al., 2018; Morillon & Baillet, 2017; Park et al., 2015). The present findings suggest that CTS in AWS is impaired especially when speech listening needs to be interwoven with speech production. In a first instance, this could be due to the fact the stuttering brain seems to be characterized by metabolic alterations (Alm, 2004), thus resulting in less stable speech-motor planning and execution (Chang & Guenther, 2020; Max et al., 2004). The listening for speaking condition entails the ability of efficiently timing the transition between listening and speaking, and appropriately plan production initiation while still attending to speech, similarly to turn-taking in conversational settings (Levinson, 2016). Given the intrinsic inefficiency in timing and implementing these very same processes for speech production in developmental stuttering (Alm, 2004, 2021b; Busan, 2020; Chang & Guenther, 2020), CTS is also affected as a result.

At the source level, in the listening-for-speaking task, we observed this reduction in the left inferior parietal cortex and in the right premotor and supplementary motor regions. All these regions are key cortical substrates for speech-motor coordination. The inferior parietal lobule (IPL), comprising the supramarginal gyrus and the angular gyrus, has been associated with a variety of functions, including verbal working memory, auditory spatial localization, sensorimotor integration, semantic processing and action-motor control (Binder et al., 2009; Binkofski & Buccino, 2018; Bzdok et al., 2016; Shum et al., 2011). Importantly, this region and the partially overlapping (non-anatomically defined) temporo-parietal junction (TPJ; Igelström & Graziano, 2017) are nodes in many dual-route models of speech and auditory processing (Friederici, 2012; Hickok et al., 2011; Hickok & Poeppel, 2004, 2007; Rauschecker, 2012). For instance, in Hickock and Poeppel's model the Sylvian Parietal Temporal (Spt) area (located between the inferior parietal lobule and the posterior part of the superior temporal gyrus, thus situated within the TPJ) is proposed to be an interface between auditory codes and motor programs for the successful sensorimotor integration during speech production, instantiated in the dorsal pathway. The IPL is also key in neurocomputational models of speech production such as the DIVA/GODIVA models (Guenther, 2016), which proposes that somatosensory error maps are stored in the IPL, where the difference between intended and actual somatic states are computed during speech production. Importantly, in the adjacent posterior STG/TPJ, auditory error maps are computed by comparing auditory feedback and predicted targets via motor efference copies (Guenther, 2016). Interestingly, even if not properly part of the "classical" cortico-basal-thalamo-cortical network involved in DS (Alm, 2004; Busan, 2020; Chang & Guenther, 2020; Craig-McQuaide et al., 2014), it is not uncommon for this region to be highlighted as part of a defective system in the brain of people who stutter (Busan et al., 2019; Neef et al., 2015; Yang et al., 2016).

On the other hand, the premotor ROI in our study encompasses the premotor cortex and the supplementary motor complex (supplementary motor area – SMA – and pre-SMA), which are essential for coordinating the composition and timing the execution of speech-motor command sequences (Alario et al., 2006; Ghosh et al., 2008; Guenther, 2016). As a matter of fact, especially in the right hemisphere, this region has been described in the DIVA/GODIVA models as a component of a feedback/control network (Guenther, 2016; Tourville & Guenther, 2011; see Chang & Guenther, 2020; Civier et al., 2013 for a perspective on DS). Compatibly, it has also been linked to mediate motor-sounds representations in auditory prediction and speech imagery (Lima et al., 2016). As foreseen by CTS functionalities, this may result in exchanging information with auditory/sensory regions, thus having the possible role of facilitating speech predictability and/or processes related to speech production. Importantly, premotor and supplementary motor regions are among the regions that are found to be most dysfunctional in DS (Busan et al., 2019; Busan, 2020; Chang & Guenther, 2020; Civier et al., 2013; A. C. Etchell et al., 2018).

Given this picture, our results suggest that when upcoming speech is required and when these structures are inherently inefficient or hinder the proper function of the neural circuit in which they are recruited, as is the case of DS, such structures cannot properly contribute to tracking the speech signal. The fact that we identified regions that are commonly associated with a dorsal stream of speech processing strengthen the connection between speech-motor abilities and auditory processing, specifically the alignment to quasi-rhythmic energy alternation in the auditory stream via bidirectional motor-auditory mapping (Friederici, 2012; Hickok & Poeppel, 2007).

The fact that we found these differences in the theta range is also interesting. The theta rhythm has often been associated with syllabic grouping across languages, more specifically to the acoustic energy fluctuations of speech sound clusters organized around an energy peak (usually a vowel; Strauß & Schwartz, 2017; see also Molinaro & Lizarazu, 2018). Indeed, in our audio stimuli, a peak was present in the theta range (3-5 Hz), reflecting such acoustic property (see Supplementary Figure 2). The syllable has been proposed to be an "interface" between the perceptual and the articulatory systems (Poeppel & Assaneo, 2020; Strauß & Schwartz, 2017). Articulators are biomechanically constrained with the possible configurations they can produce and at which speed they can be executed; the syllable represents the optimal motor-programming unit that the neural system can send for execution to the motor system (Guenther, 2016; Poeppel & Assaneo, 2020). Crucially, the ability to automatically activate syllabic motor units associated to learned sound sequences via the basal ganglia motor loop connected to the pre-SMA and SMA seems to be impaired in DS (Alm, 2004, 2021a, 2021b; Busan, 2020; Chang & Guenther, 2020; Civier et al., 2013).

Therefore, it appears that there is a circular relationship that from biomechanical articulation constraints, via neural motor program units, lead to the acoustic - and hence perceptual - phenomenon of syllabic rhythm tracking (Poeppel & Assaneo, 2020; Strauß & Schwartz, 2017). We believe that this proposed circle of joint causes is strictly related to the results obtained in the present study: the speech-motor production system is involved in tracking acoustic properties that themselves arise from articulatory-motor constraints; when such a system is unstable, perceptual tracking is less efficient as well, more noticeable when listening and speaking are interwoven and partially overlapping pools of resources are required, thus overburdening an already unstable system. From a behavioral point of view, the presence of slower RTs in AWS supports this vision.

Weaker inter-hemispheric connectivity among bilateral auditory and sensorimotor regions in developmental stuttering when listening to speech

We found group differences in the connectivity patterns in both listening conditions. During the listening-only task, we found weaker directional connectivity from the left STG (auditory regions) to the right IFG and from the right IFG toward the left IPL/TPJ. In the listening-for-speaking task, we found weaker directional connectivity from the right STG to the left primary sensorimotor (SM) regions and to the left IPL/TPJ. While we do not interpret the different patterns across listening conditions, these results likely reflect lower availability of neural resources for information exchange between recruited regions that are instrumental for auditory and motor processing and integration. More specifically, evidence that AWS display reduced connectivity (when compared to TFA) is compatible with recent proposals suggesting the presence of a general metabolic deficit in the stuttering brain (Alm, 2021a; see also Busan et al., 2019; Han et al., 2019; Maguire et al., 2021; Turk et al., 2021).

Importantly, in the listening-for-speaking task, we found a significant negative correlation between RTs and strength of right STG → left SM cortex connectivity: faster RTs are associated with increased connectivity between these regions. This may indicate that efficiently sending rhythmic auditory information to the primary sensorimotor cortex when speech listening has to be managed with (overt) upcoming speech-motor engagement leads to more ready speech production, possibly as a result of more efficient CTS and smoother transitioning between listening and speaking with concomitant speech planning. Note that this correlation, when explored separately for AWS and TFA, is not strongly evident in AWS ($r = -0.44$, $p = 0.15$) but is present in TFA ($r = -0.58$, $p = 0.045$; see Supplementary Table 1). As a further indication, we would also like to highlight that, albeit statistically not significant, an interesting trend is present in AWS when looking at this very same connectivity pattern and SSI-4, where a negative relation is found ($r = -0.55$, $p = 0.06$; see Supplementary Table 1): higher SSI-4 scores - hence, more severe stuttering - are associated with weaker right STG → left SM cortex connectivity (Supplementary Figure 3). These pieces together strongly suggest that stuttering may be associated with weaker connectivity between auditory and primary sensorimotor regions at a frequency range fundamental for cortical tracking of syllabic units in the theta range, in turn leading to poorer behavioral performance in response times. This may be compatible with proposals according to which an effective connection between these regions may be helpful for better managing (or “by-passing”) dysfluencies, perhaps by exploiting rhythmic or tracked cues (A. C. Etchell et al., 2014).

Consistently with these suggestive although marginal evidence, it is important to note that right-hemispheric processes, especially in fronto-temporal regions, are often reported as neural markers of stuttering trait and state (e.g., Belyk et al., 2015, 2017; Brown et al., 2005a; Budde et al., 2014; Craig-McQuaide et al., 2014; A. C. Etchell et al., 2014, 2018; Ingham et al., 2012; Neef et al., 2015), suggesting that they may have a role in compensatory (as well as in pathological) speech-motor programming and execution processes in AWS (Busan et al., 2019; A. C. Etchell et al., 2014; Neef et al., 2015, 2016, 2018b, 2023).

Significance of present outcomes for CTS and DS research

The present findings may advance research on both CTS and DS. More specifically, they suggest that 1) CTS needs neural resources that sustain sensorimotor processes for facilitating speech perception and intelligibility, 2) DS may lead to suboptimal CTS, especially when additional resources are needed for supporting concomitant speech preparation for upcoming production, and 3) DS is a more complex neurodevelopmental disorder, i.e., not only impairing speech programming and production. Further research should clarify the extent to which DS impacts CTS (and viceversa), and what the possible effects in people's everyday life may be, hence widening the scope of possible intervention for stuttering. This is especially important also at the light of recent evidence suggesting that auditory-motor coupling (and individual speech production rates) may explain performances in speech comprehension tasks (Lubinus et al., 2023). Speculatively, less efficient CTS may be related with more effortful spoken language comprehension at a subtle level. Compatibly, the results reported in Gastaldon et al. (2023) goes in this direction: AWS seem less efficient in generating predictions during listening, hypothesized as a result of the inability to fully exploit their speech-motor network. Thus, further studies should investigate whether there is a causal link between CTS and specific processes of speech comprehension such as prediction, and how this causal chain may impact people with different speech and language deficits, especially in interactive contexts. In conclusion, research should move towards turn-based and conversational contexts (e.g., Jackson et al., 2021; Weiss, 1995) to also investigate possible subtle differences in how spoken language comprehension is attained in the stuttering brain.

Conclusions

The present work suggests that CTS recruits (pre-)motor regions and regions responsible for sensorimotor integration, as well as auditory regions, supporting the view proposing an interaction between these networks also in speech/language perception (Pickering & Garrod, 2013; Skipper et al., 2017), in addition to be instrumental in orchestrating successful speech production (Guenther, 2016; Hickok et al., 2011). This process seems to work less efficiently in DS, especially when additional neural resources are needed for managing listening-for-speaking conditions, as usually happens in more ecological communicative situations (Neef & Chang, 2024). A better understanding of CTS processes in DS under various circumstances may be informative for improving rehabilitation solutions for stuttering.

Contribution statement: SG, NM and ML conceptualized the study. SG and ML performed analyses. SG, PB and ML wrote the initial draft (Introduction: SG, PB; Materials and Methods: SG, ML; Results: SG, ML; Discussion: SG, PB; Conclusion: SG, PB). All authors provided feedback and revisions and contributed to the interpretation of the results.

Conflicts of interest: None.

References

Alario, F. X., Chainay, H., Lehericy, S., & Cohen, L. (2006). The role of the supplementary motor area (SMA) in word production. *Brain Research*, *1076*(1), 129–143. https://doi.org/10.1016/j.brainres.2005.11.104

Alm, P. A. (2004). Stuttering and the basal ganglia circuits: A critical review of possible relations. *Journal of Communication Disorders*, *37*(4), 325–369. https://doi.org/10.1016/j.jcomdis.2004.03.001

Alm, P. A. (2021a). Stuttering: A Disorder of Energy Supply to Neurons? *Frontiers in Human Neuroscience*, *0*, 289. https://doi.org/10.3389/FNHUM.2021.662204

Alm, P. A. (2021b). The Dopamine System and Automatization of Movement Sequences: A Review With Relevance for Speech and Stuttering. *Frontiers in Human Neuroscience*, *15*, 663. https://doi.org/10.3389/FNHUM.2021.661880/BIBTEX

American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders: DSM-5™, 5th ed. In *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

Assaneo, M. F., & Poeppel, D. (2018). The coupling between auditory and motor cortices is rate-restricted: Evidence for an intrinsic speech-motor rhythm. *Science Advances*, *4*(2), eaao3842. https://doi.org/10.1126/sciadv.aao3842

Assaneo, M. F., Ripollés, P., Tichenor, S. E., Yaruss, J. S., & Jackson, E. S. (2022). The Relationship Between Auditory-Motor Integration, Interoceptive Awareness, and Self-Reported Stuttering Severity. *Frontiers in Integrative Neuroscience*, *16*(May), 1–11. https://doi.org/10.3389/fnint.2022.869571

Barnes, T. D., Wozniak, D. F., Gutierrez, J., Han, T. U., Drayna, D., & Holy, T. E. (2016). A Mutation Associated with Stuttering Alters Mouse Pup Ultrasonic Vocalizations. *Current Biology*, *26*(8), 1009–1018. https://doi.org/10.1016/J.CUB.2016.02.068

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Belyk, M., Kraft, S. J., & Brown, S. (2015). Stuttering as a trait or state - an ALE meta-analysis of neuroimaging studies. *European Journal of Neuroscience*, *41*(2), 275–284. https://doi.org/10.1111/ejn.12765

Belyk, M., Kraft, S. J., & Brown, S. (2017). Stuttering as a trait or a state revisited: motor system involvement in persistent developmental stuttering. *European Journal of Neuroscience*, *45*(4), 622–624. https://doi.org/10.1111/EJN.13512

Benito-Aragón, C., Gonzalez-Sarmiento, R., Liddell, T., Diez, I., d'Oleire Uquillas, F., Ortiz-Terán, L., Bueichekú, E., Chow, H. M., Chang, S. E., & Sepulcre, J. (2020). Neurofilament-lysosomal genetic intersections in the cortical network of stuttering. *Progress in Neurobiology*, *184*, 101718. https://doi.org/10.1016/j.pneurobio.2019.101718

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. *Cerebral Cortex*, *19*(12), 2767–2796. https://doi.org/10.1093/CERCOR/BHP055

Binkofski, F., & Buccino, G. (2018). The role of the parietal cortex in sensorimotor transformations and action coding. In *Handbook of Clinical Neurology* (1st ed., Vol. 151). Elsevier B.V. https://doi.org/10.1016/B978-0-444-63622-5.00024-3

Brown, S., Ingham, R. J., Ingham, J. C., Laird, A. R., & Fox, P. T. (2005a). Stuttered and fluent speech production: An ALE meta-analysis of functional neuroimaging studies. *Human Brain Mapping*, *25*(1), 105–117. https://doi.org/10.1002/hbm.20140

Brown, S., Ingham, R. J., Ingham, J. C., Laird, A. R., & Fox, P. T. (2005b). Stuttered and fluent speech production: An ALE meta-analysis of functional neuroimaging studies. *Human Brain Mapping*, *25*(1), 105–117. https://doi.org/10.1002/HBM.20140

Budde, K. S., Barron, D. S., & Fox, P. T. (2014). Stuttering, induced fluency, and natural fluency: A hierarchical series of activation likelihood estimation meta-analyses. *Brain and Language*, *139*, 99–107. https://doi.org/10.1016/J.BANDL.2014.10.002

Busan, P. (2020). Developmental stuttering and the role of the supplementary motor cortex. *Journal of Fluency Disorders*, 105763. https://doi.org/10.1016/j.jfludis.2020.105763

Busan, P., Del Ben, G., Russo, L. R., Bernardini, S., Natarelli, G., Arcara, G., Manganotti, P., & Battaglini, P. P. (2019). Stuttering as a matter of delay in neural activation: A combined TMS/EEG study. *Clinical Neurophysiology*, *130*(1), 61–76. https://doi.org/10.1016/J.CLINPH.2018.10.005

Bzdok, D., Hartwigsen, G., Reid, A., Laird, A. R., Fox, P. T., & Eickhoff, S. B. (2016). Left inferior parietal lobe engagement in social cognition and language. In *Neuroscience and Biobehavioral Reviews* (Vol. 68, pp. 319–334). Elsevier Ltd. https://doi.org/10.1016/j.neubiorev.2016.02.024

Chang, S. E., Chow, H. M., Wieland, E. A., & McAuley, J. D. (2016). Relation between functional connectivity and rhythm discrimination in children who do and do not stutter. *NeuroImage: Clinical*, *12*, 442–450. https://doi.org/10.1016/J.NICL.2016.08.021

Chang, S. E., Garnett, E. O., Etchell, A., & Chow, H. M. (2019). Functional and Neuroanatomical Bases of Developmental Stuttering: Current Insights. *The Neuroscientist*, *25*(6), 566–582. https://doi.org/10.1177/1073858418803594

Chang, S. E., & Guenther, F. H. (2020). Involvement of the Cortico-Basal Ganglia-Thalamocortical Loop in Developmental Stuttering. *Frontiers in Psychology*, *10*, 3088. https://doi.org/10.3389/FPSYG.2019.03088/BIBTEX

Chow, H. M., Garnett, E. O., Li, H., Etchell, A., Sepulcre, J., Drayna, D., Chugani, D. C., & Chang, S.-E. (2020). Linking lysosomal enzyme targeting genes and energy metabolism with altered gray matter volume in children with persistent stuttering. *Neurobiology of Language*, 1–40. https://doi.org/10.1162/nol_a_00017

Civier, O., Bullock, D., Max, L., & Guenther, F. H. (2013). Computational modeling of stuttering caused by impairments in a basal ganglia thalamo-cortical circuit involved in syllable selection and initiation. *Brain and Language*, *126*(3), 263–278. https://doi.org/10.1016/j.bandl.2013.05.016

Corps, R. E., Gambi, C., & Pickering, M. J. (2018). Coordinating Utterances During Turn-Taking: The Role of Prediction, Response Preparation, and Articulation. *Discourse Processes*, *55*(2), 230–240. https://doi.org/10.1080/0163853X.2017.1330031

Craig-McQuaide, A., Akram, H., Zrinzo, L., & Tripoliti, E. (2014). A review of brain circuitries involved in stuttering. *Frontiers in Human Neuroscience*, *8*(November), 1–20. https://doi.org/10.3389/fnhum.2014.00884

Craig, A., Blumgart, E., & Tran, Y. (2009). The impact of stuttering on the quality of life in adults who stutter. *Journal of Fluency Disorders*, *34*(2), 61–71. https://doi.org/10.1016/J.JFLUDIS.2009.05.002

Daliri, A., & Max, L. (2015). Modulation of auditory processing during speech movement planning is limited in adults who stutter. *Brain and Language*, *143*, 59–68. https://doi.org/10.1016/J.BANDL.2015.03.002

Daliri, A., & Max, L. (2018). Stuttering adults' lack of pre-speech auditory modulation normalizes when speaking with delayed auditory feedback. *Cortex*, *99*, 55–68. https://doi.org/10.1016/J.CORTEX.2017.10.019

Etchell, A., Adhikari, A., Weinberg, L. S., Choo, A. L., Garnett, E. O., Chow, H. M., & Chang, S.-E. (2018). A systematic literature review of sex differences in childhood language and brain development. *Neuropsychologia*, *114*, 19–31. https://doi.org/10.1016/j.neuropsychologia.2018.04.011

Etchell, A. C., Civier, O., Ballard, K. J., & Sowman, P. F. (2018). A systematic literature review of neuroimaging research on developmental stuttering between 1995 and 2016. *Journal of Fluency Disorders*, *55*, 6–45. https://doi.org/10.1016/j.jfludis.2017.03.007

Etchell, A. C., Johnson, B. W., & Sowman, P. F. (2014). Behavioral and multimodal neuroimaging evidence for a deficit in brain timing networks in stuttering: A hypothesis and theory. *Frontiers in Human Neuroscience*, *8*(JUNE), 467. https://doi.org/10.3389/fnhum.2014.00467

Etchell, A. C., Ryan, M., Martin, E., Johnson, B. W., & Sowman, P. F. (2016). Abnormal time course of low beta modulation in non-fluent preschool children: A magnetoencephalographic study of rhythm tracking. *NeuroImage*, *125*, 953–963. https://doi.org/10.1016/j.neuroimage.2015.10.086

Falk, S., Müller, T., & Dalla Bella, S. (2015). Non-verbal sensorimotor timing deficits in children and adolescents who stutter. *Frontiers in Psychology*, *6*, 847. https://doi.org/10.3389/fpsyg.2015.00847

Friederici, A. D. (2012). The cortical language circuit: from auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, *16*(5), 262–268. https://doi.org/10.1016/J.TICS.2012.04.001

Frigerio-Domingues, C., & Drayna, D. (2017). Genetic contributions to stuttering: the current evidence. *Molecular Genetics & Genomic Medicine*, *5*(2), 95–102. https://doi.org/10.1002/MGG3.276

Gastaldon, S., Arcara, G., Navarrete, E., & Peressotti, F. (2020). Commonalities in alpha and beta neural desynchronizations during prediction in language comprehension and production. *Cortex*, *133*, 328–345. https://doi.org/10.1016/j.cortex.2020.09.026

Gastaldon, S., Busan, P., Arcara, G., & Peressotti, F. (2023). Inefficient speech-motor control affects predictive speech comprehension: atypical electrophysiological correlates in stuttering. *Cerebral Cortex*, *33*(11), 6834–6851. https://doi.org/10.1093/cercor/bhad004

Ghaderi, A. H., Andevari, M. N., & Sowman, P. F. (2018). Evidence for a Resting State Network Abnormality in Adults Who Stutter. *Frontiers in Integrative Neuroscience*, *12*, 16. https://doi.org/10.3389/fnint.2018.00016

Ghosh, S. S., Tourville, J. A., & Guenther, F. H. (2008). A neuroimaging study of premotor lateralization and cerebellar involvement in the production of phonemes and syllables. *Journal of Speech, Language, and Hearing Research*, *51*(5), 1183–1202. https://doi.org/10.1044/1092-4388(2008/07-0119)

Goslin, J., Galluzzi, C., & Romani, C. (2014). PhonItalia: a phonological lexicon for Italian. *Behavior Research Methods*, *46*(3), 872–886. https://doi.org/10.3758/s13428-013-0400-8

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, *3*(2), 424–438.

Gross, J., Kujala, J., Hämäläinen, M., Timmermann, L., Schnitzler, A., & Salmelin, R. (2001). Dynamic imaging of coherent sources: Studying neural interactions in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(2), 694–699. https://doi.org/10.1073/PNAS.98.2.694/ASSET/5DFC00E6-B2AF-4B10-BCF0-A0D9FA0535C7/ASSETS/GRAPHIC/PQ0215386004.JPEG

Guenther, F. H. (2016). *Neural Control of Speech*. The MIT Press. https://doi.org/10.7551/mitpress/10471.001.0001

Halag-Milo, T., Stoppelman, N., Kronfeld-Duenias, V., Civier, O., Amir, O., Ezrati-Vinacour, R., & Ben-Shachar, M. (2016). Beyond production: Brain responses during speech perception in adults who stutter. *NeuroImage: Clinical*, *11*, 328–338. https://doi.org/10.1016/j.nicl.2016.02.017

Han, T. U., Root, J., Reyes, L. D., Huchinson, E. B., Du Hoffmann, J., Lee, W. S., Barnes, T. D., & Drayna, D. (2019). Human GNPTAB stuttering mutations engineered into mice cause vocalization deficits and astrocyte pathology in the corpus callosum. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(35), 17515–17524. https://doi.org/10.1073/PNAS.1901480116/SUPPL_FILE/PNAS.1901480116.SAPP.PDF

Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor Integration in Speech Processing: Computational Basis and Neural Organization. *Neuron*, *69*(3), 407–422. https://doi.org/10.1016/j.neuron.2011.01.019

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, *92*(1–2), 67–99.

https://doi.org/10.1016/j.cognition.2003.10.011

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

Igelström, K. M., & Graziano, M. S. A. (2017). The inferior parietal lobule and temporoparietal junction: A network perspective. *Neuropsychologia*, *105*, 70–83. https://doi.org/10.1016/j.neuropsychologia.2017.01.001

Ingham, R. J., Grafton, S. T., Bothe, A. K., & Ingham, J. C. (2012). Brain activity in adults who stutter: Similarities across speaking tasks and correlations with stuttering frequency and speaking rate. *Brain and Language*, *122*(1), 11–24. https://doi.org/10.1016/j.bandl.2012.04.002

Jackson, E. S., Miller, L. R., Warner, H. J., & Yaruss, J. S. (2021). Adults who stutter do not stutter during private speech. *Journal of Fluency Disorders*, *70*, 105878. https://doi.org/10.1016/J.JFLUDIS.2021.105878

Jenson, D., Bowers, A. L., Hudock, D., & Saltuklaroglu, T. (2020). The Application of EEG Mu Rhythm Measures to Neurophysiological Research in Stuttering. *Frontiers in Human Neuroscience*, *13*, 458. https://doi.org/10.3389/fnhum.2019.00458

Jenson, D., Reilly, K. J., Harkrider, A. W., Thornton, D., & Saltuklaroglu, T. (2018). Trait related sensorimotor deficits in people who stutter: An EEG investigation of µ rhythm dynamics during spontaneous fluency. *NeuroImage: Clinical*, *19*, 690–702. https://doi.org/10.1016/J.NICL.2018.05.026

Joos, K., Ridder, D. De, Boey, R. A., & Vanneste, S. (2014). Functional connectivity changes in adults with developmental stuttering: A preliminary study using quantitative electro-encephalography. *Frontiers in Human Neuroscience*, *8*(OCT). https://doi.org/10.3389/fnhum.2014.00783

Kang, C., & Drayna, D. (2012). A role for inherited metabolic deficits in persistent developmental stuttering. *Molecular Genetics and Metabolism*, *107*(3), 276–280. https://doi.org/10.1016/J.YMGME.2012.07.020

Kang, C., Riazuddin, S., Mundorff, J., Krasnewich, D., Friedman, P., Mullikin, J. C., & Drayna, D. (2010). Mutations in the Lysosomal Enzyme–Targeting Pathway and Persistent Stuttering. *New England Journal of Medicine*, *362*(8), 677–685. https://doi.org/10.1056/NEJMOA0902630/SUPPL_FILE/NEJM_KANG_677SA1.PDF

Keitel, A., Gross, J., & Kayser, C. (2018). Perceptually relevant speech tracking in auditory and motor cortex reflects distinct linguistic features. *PLOS Biology*, *16*(3), e2004473. https://doi.org/10.1371/journal.pbio.2004473

Kim, K. S., Daliri, A., Flanagan, J. R., & Max, L. (2020). Dissociated Development of Speech and Limb Sensorimotor Learning in Stuttering: Speech Auditory-motor Learning is Impaired in Both Children and Adults Who Stutter. *Neuroscience*, *451*, 1–21. https://doi.org/10.1016/J.NEUROSCIENCE.2020.10.014

Kraft, S. J., & Yairi, E. (2011). Genetic Bases of Stuttering: The State of the Art, 2011. *Folia Phoniatrica et Logopaedica*, *64*(1), 34–47. https://doi.org/10.1159/000331073

Kronfeld-Duenias, V., Amir, O., Ezrati-Vinacour, R., Civier, O., & Ben-Shachar, M. (2016). Dorsal and ventral language pathways in persistent developmental stuttering. *Cortex*, *81*, 79–92. https://doi.org/10.1016/j.cortex.2016.04.001

Ladányi, E., Persici, V., Fiveash, A., Tillmann, B., & Gordon, R. L. (2020). Is atypical rhythm a risk factor for developmental speech and language disorders? *Wiley Interdisciplinary Reviews: Cognitive Science*, *March*, e1528. https://doi.org/10.1002/wcs.1528

Levinson, S. C. (2016). Turn-taking in Human Communication - Origins and Implications for Language Processing. *Trends in Cognitive Sciences*, *20*(1), 6–14. https://doi.org/10.1016/j.tics.2015.10.010

Lima, C. F., Krishnan, S., & Scott, S. K. (2016). Roles of Supplementary Motor Areas in Auditory Processing and Auditory Imagery. In *Trends in Neurosciences* (Vol. 39, Issue 8, pp. 527–542). Elsevier Ltd. https://doi.org/10.1016/j.tins.2016.06.003

Lizarazu, M., Lallier, M., Molinaro, N., Bourguignon, M., Paz-Alonso, P. M., Lerma-Usabiaga, G., & Carreiras, M. (2015). Developmental evaluation of atypical auditory sampling in dyslexia: Functional and structural evidence. *Human Brain Mapping*, *36*(12), 4986–5002. https://doi.org/10.1002/HBM.22986

Lizarazu, M., Scotto di Covella, L., van Wassenhove, V., Rivière, D., Mizzi, R., Lehongre, K., Hertz-Pannier, L., & Ramus, F. (2021). Neural entrainment to speech and nonspeech in dyslexia: Conceptual replication and extension of previous investigations. *Cortex*, *137*, 160–178. https://doi.org/10.1016/j.cortex.2020.12.024

Lubinus, C., Keitel, A., Obleser, J., Poeppel, D., & Rimmele, J. M. (2023). Explaining flexible continuous speech comprehension from individual motor rhythms. *Proceedings of the Royal Society B*, *290*(1994). https://doi.org/10.1098/RSPB.2022.2410

Maguire, G. A., Yoo, B. R., & SheikhBahaei, S. (2021). Investigation of Risperidone Treatment Associated With Enhanced Brain Activity in Patients Who Stutter. *Frontiers in Neuroscience*, *15*, 598949. https://doi.org/10.3389/FNINS.2021.598949/BIBTEX

Max, L., Guenther, F. H., Gracco, V. L., Ghosh, S. S., & Wallace, M. E. (2004). Unstable or Insufficiently Activated Internal Models and Feedback-Biased Motor Control as Sources of Dysfluency: A Theoretical Model of Stuttering. *Contemporary Issues in Communication Science and Disorders*, *31*(Spring), 105–122. https://doi.org/10.1044/cicsd_31_S_105

Molinaro, N., & Lizarazu, M. (2018). Delta(but not theta)-band cortical entrainment involves speech-specific processing. *European Journal of Neuroscience*, *48*(7), 2642–2650. https://doi.org/10.1111/ejn.13811

Molinaro, N., Lizarazu, M., Lallier, M., Bourguignon, M., & Carreiras, M. (2016). Out-of-synchrony speech entrainment in developmental dyslexia. *Human Brain Mapping*, *37*(8), 2767–2783. https://doi.org/10.1002/hbm.23206

Morillon, B., & Baillet, S. (2017). Motor origin of temporal predictions in auditory attention. *Proceedings of the National Academy of Sciences*, *114*(42), E8913–E8921. https://doi.org/10.1073/pnas.1705373114

Nang, C., Hersh, D., Milton, K., & Lau, S. R. (2018). The Impact of Stuttering on Development of Self-Identity, Relationships, and Quality of Life in Women Who Stutter. *American Journal of Speech-Language Pathology*, *27*(3S), 1244–1258. https://doi.org/10.1044/2018_AJSLP-ODC11-17-0201

Neef, N. E., Angstadt, M., Koenraads, S. P. C., & Chang, S. E. (2023). Dissecting structural connectivity of the left and right inferior frontal cortex in children who stutter. *Cerebral Cortex*, *33*(7), 4085–4100. https://doi.org/10.1093/CERCOR/BHAC328

Neef, N. E., Anwander, A., Bütfering, C., Schmidt-Samoa, C., Friederici, A. D., Paulus, W., & Sommer, M. (2018a). Structural connectivity of right frontal hyperactive areas scales with stuttering severity. *Brain*, *141*(1), 191–204. https://doi.org/10.1093/brain/awx316

Neef, N. E., Anwander, A., Bütfering, C., Schmidt-Samoa, C., Friederici, A. D., Paulus, W., & Sommer, M. (2018b). Structural connectivity of right frontal hyperactive areas scales with stuttering severity. *Brain*, *141*(1), 191–204. https://doi.org/10.1093/brain/awx316

Neef, N. E., Anwander, A., & Friederici, A. D. (2015). The Neurobiological Grounding of Persistent Stuttering: from Structure to Function. In *Current Neurology and Neuroscience Reports* (Vol. 15, Issue 9, pp. 1–11). Current Medicine Group LLC 1. https://doi.org/10.1007/s11910-015-0579-4

Neef, N. E., Bütfering, C., Anwander, A., Friederici, A. D., Paulus, W., & Sommer, M. (2016). Left posterior-dorsal area 44 couples with parietal areas to promote speech fluency, while right area 44 activity promotes the stopping of motor responses. *NeuroImage*, *142*, 628–644. https://doi.org/10.1016/J.NEUROIMAGE.2016.08.030

Neef, N. E., & Chang, S.-E. (2024). Knowns and unknowns about the neurobiology of stuttering. *PLOS Biology*, *22*(2), e3002492. https://doi.org/10.1371/journal.pbio.3002492

Neef, N. E., Korzeczek, A., Primaßin, A., Wolff von Gudenberg, A., Dechent, P., Riedel, C. H., Paulus, W., & Sommer, M. (2022). White matter tract strength correlates with therapy outcome in persistent developmental stuttering. *Human Brain Mapping*, *43*(11), 3357–3374. https://doi.org/10.1002/HBM.25853

Nuwer, M. R., Comi, G., Emerson, R., Fuglsang-Frederiksen, A., Guérit, J.-M., Hinrichs, H., Ikeda, A., Jose C. Luccas, F., & Rappelsburger, P. (1998). IFCN standards for digital recording of clinical EEG. *Electroencephalography and Clinical Neurophysiology*, *106*(3), 259–261. https://doi.org/10.1016/S0013-4694(97)00106-5

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*. https://doi.org/10.1155/2011/156869

Park, H., Ince, R. A. A., Schyns, P. G., Thut, G., & Gross, J. (2015). Frontal Top-Down Signals Increase Coupling of Auditory Low-Frequency Oscillations to Continuous Speech in Human Listeners. *Current Biology*, *25*(12), 1649–1653. https://doi.org/10.1016/J.CUB.2015.04.049

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(4), 329–347. https://doi.org/10.1017/S0140525X12001495

Poeppel, D., & Assaneo, M. F. (2020). Speech rhythms and their neural foundations. *Nature Reviews Neuroscience*, 1–13. https://doi.org/10.1038/s41583-020-0304-4

Poeppel, D., & Teng, X. (2020). Entrainment in Human Auditory Cortex: Mechanism and Functions. In *Reference Module in Neuroscience and Biobehavioral Psychology*. Elsevier. https://doi.org/10.1016/B978-0-12-809324-5.00018-3

Rauschecker, J. P. (2012). Ventral and dorsal streams in the evolution of speech and language. *Frontiers in Evolutionary Neuroscience*, *4*(MAY), 18154. https://doi.org/10.3389/FNEVO.2012.00007/BIBTEX

Riley, G. D. (2009). *The stuttering severity instrument for adults and children (SSI-4). 4th Edition. PRO-ED Inc.*

Rimmele, J. M., Morillon, B., Poeppel, D., & Arnal, L. H. (2018). Proactive Sensing of Periodic and Aperiodic Auditory Patterns. *Trends in Cognitive Sciences*, *22*(10), 870–882. https://doi.org/10.1016/J.TICS.2018.08.003

Roux, F., Armstrong, B. C., & Carreiras, M. (2017). Chronset: An automated tool for detecting speech onset. *Behavior Research Methods*, *49*(5), 1864–1881. https://doi.org/10.3758/s13428-016-0830-1

Saltuklaroglu, T., Harkrider, A. W., Thornton, D., Jenson, D., & Kittilstved, T. (2017). EEG Mu (µ) rhythm spectra and oscillatory activity differentiate stuttering from non-stuttering adults. *NeuroImage*, *153*, 232–245. https://doi.org/10.1016/j.neuroimage.2017.04.022

Schneider, T., & Neumaier, A. (2001). Algorithm 808: ARFIT - A Matlab package for the estimation of parameters and eigenmodes of multivariate autoregressive models. *ACM Transactions on Mathematical Software*, *27*(1), 58–65. https://doi.org/10.1145/382043.382316

Seth, A. K., Barrett, A. B., & Barnett, L. (2015). Granger Causality Analysis in Neuroscience and Neuroimaging. *Journal of Neuroscience*, *35*(8), 3293–3297. https://doi.org/10.1523/JNEUROSCI.4399-14.2015

Shum, M., Shiller, D. M., Baum, S. R., & Gracco, V. L. (2011). Sensorimotor integration for speech motor learning involves the inferior parietal cortex. *European Journal of Neuroscience*, *34*(11), 1817–1822. https://doi.org/10.1111/J.1460-9568.2011.07889.X

Skipper, J. I., Devlin, J. T., & Lametti, D. R. (2017). The hearing ear is always found close to the speaking tongue: Review of the role of the motor system in speech perception. *Brain and Language*, *164*, 77–105. https://doi.org/10.1016/J.BANDL.2016.10.004

Sommer, M., Koch, M. A., Paulus, W., Weiller, C., & Büchel, C. (2002). Disconnection of speech-relevant brain areas in persistent developmental stuttering. *The Lancet*, *360*(9330), 380–383. https://doi.org/10.1016/S0140-6736(02)09610-1

Strauß, A., & Schwartz, J. L. (2017). The syllable in the light of motor skills and neural oscillations. *Language, Cognition and Neuroscience*, *32*(5), 562–569. https://doi.org/10.1080/23273798.2016.1253852

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

Turk, A. Z., Lotfi Marchoubeh, M., Fritsch, I., Maguire, G. A., & SheikhBahaei, S. (2021). Dopamine, vocalization, and astrocytes. *Brain and Language*, *219*, 104970. https://doi.org/10.1016/j.bandl.2021.104970

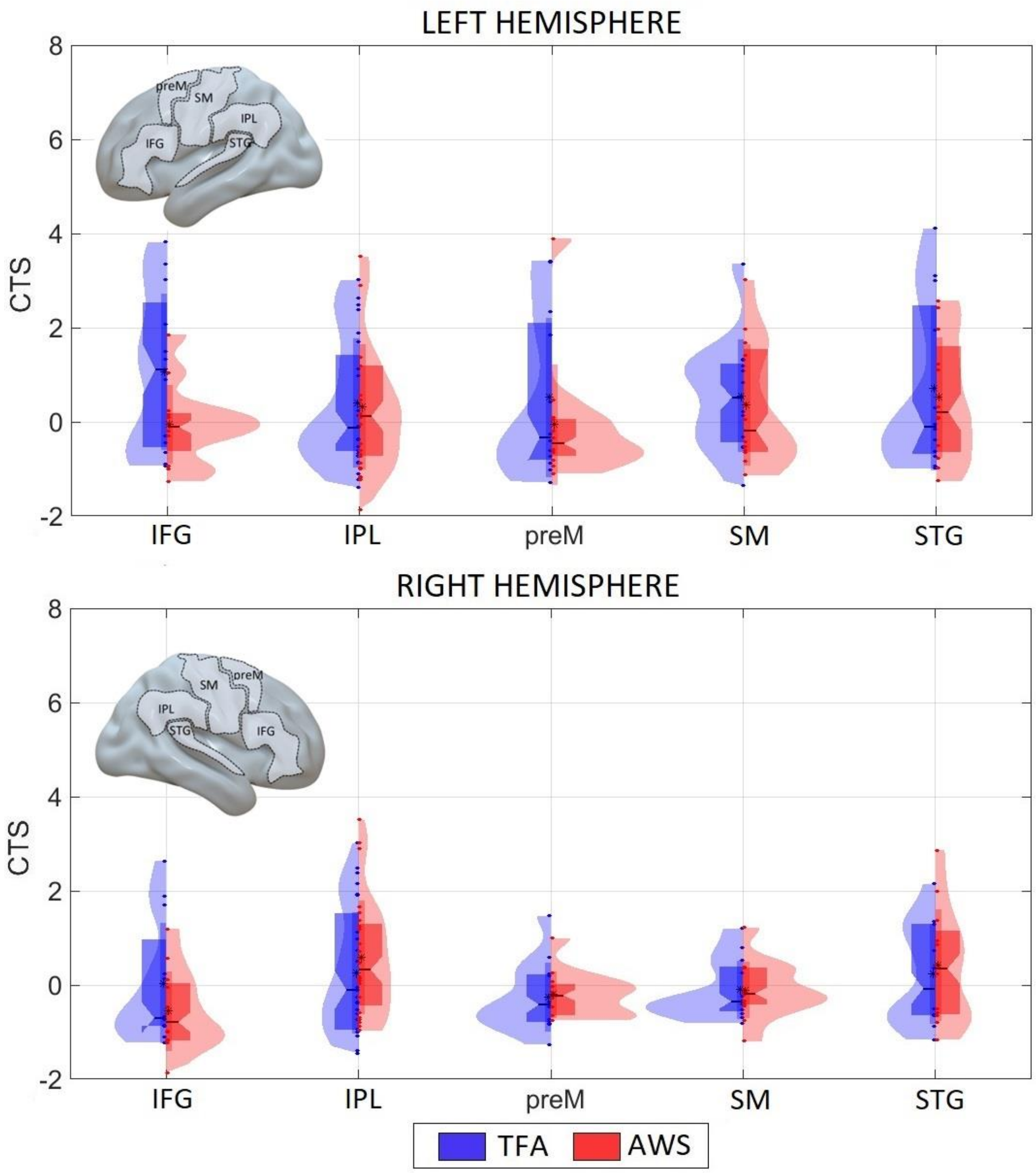
Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. (2002). Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. *NeuroImage*, *15*(1), 273–289. https://doi.org/10.1006/NIMG.2001.0978

Watkins, K. E., Smith, S. M., Davis, S., & Howell, P. (2008). Structural and functional abnormalities of the motor system in developmental stuttering. *Brain*, *131*(1), 50–59. https://doi.org/10.1093/BRAIN/AWM241

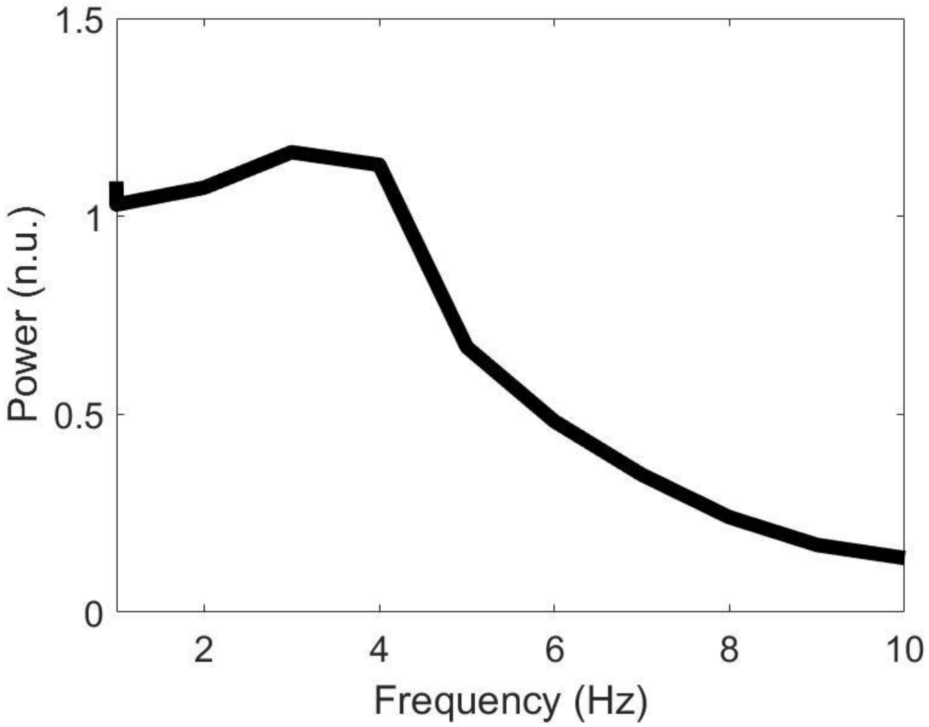
Weiss, A. L. (1995). Conversational demands and their effects on fluency and stuttering. *Topics in Language Disorders*, *15*(3), 18–31. https://doi.org/10.1097/00011363-199505000-00004

Yang, Y., Jia, F., Siok, W. T., & Tan, L. H. (2016). Altered functional connectivity in persistent developmental stuttering. *Scientific Reports*, *6*(1), 1–8. https://doi.org/10.1038/srep19128

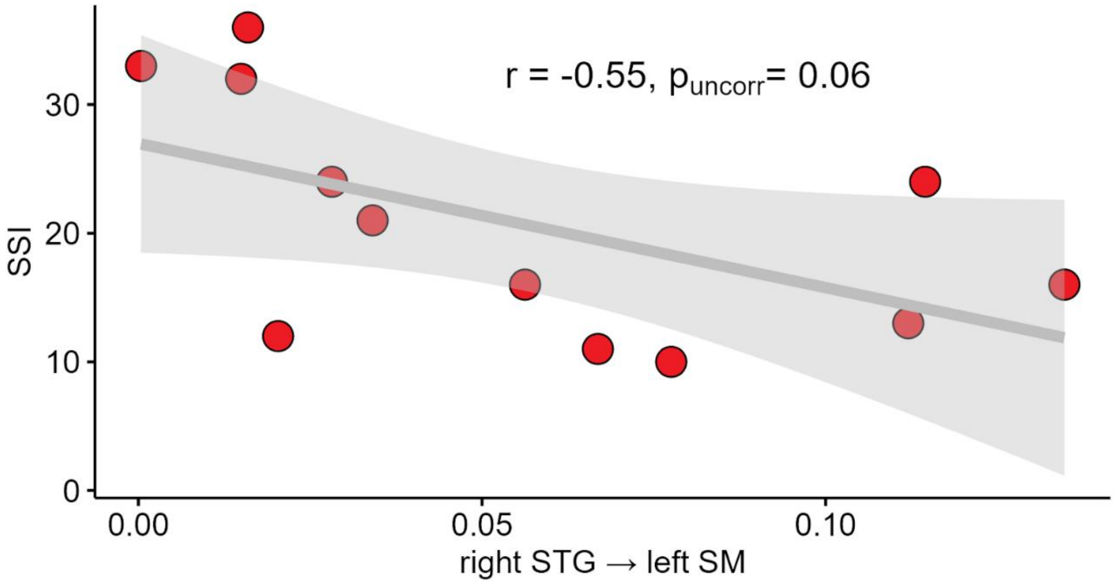
SUPPLEMENTARY MATERIAL



Supplementary Figure 1: Source-level CTS in the listening-only task. No statistically significant differences were found.



Supplementary Figure 2: Spectral characteristics of the speech envelope. We computed the power spectrum of the speech envelope for the frequencies that phase-synchronized with the electrophysiological brain activity (i.e., <10 Hz). The more prominent amplitude modulations were between 3 and 4 Hz, as expected from previous studies (Park et al., 2015; Lizarazu et al., 2019).



Supplementary Figure 3. Scatterplot of the correlation between right STG → left SM and SSI in AWS. Although statistically not significant, a negative trend is observable whereby reduced connectivity from the right STG to the left SM cortex is associated with higher SSI-4 scores (Stuttering Severity Instrument 4; Riley, 2009).

	ALL PARTICIPANTS		**AWS**				**TFA**	
	RT		**RT**		**SSI**		**RT**	
	r	**p**	**r**	**p**	**r**	**p**	**r**	**p**
SENSOR_COH_LEFT	-0.120	0.576	0.101	0.756	-0.504	0.095	0.068	0.834
SENSOR_COH_RIGHT	-0.155	0.470	-0.140	0.664	-0.508	0.092	0.177	0.582
SOURCE_COH_L_IPL	-0.038	0.859	-0.051	0.876	-0.272	0.392	0.458	0.134
SOURCE_COH_R_PREM	-0.083	0.699	0.136	0.674	-0.386	0.215	0.226	0.479
CONN_R_STG_L_IPL	-0.199	0.350	0.072	0.824	-0.444	0.149	-0.154	0.632
CONN_R_STG_L_SM	**-0.556**	**0.005**	-0.439	0.153	-0.551	0.064	**-0.576**	**0.050**

Supplementary Table 1: Correlations between the neural measures found to be different across groups in the listening-for-speaking task (coherence at sensor and source level, connectivity at theta) and behavioral measures (response times and SSI-4). Separate correlations for AWS and TFA are reported only as exploratory.