

Research Article

Cortical Tracking of Speech Is Reduced in Adults Who Stutter When Listening for Speaking

Simone Gastaldon,^{a,b}  Pierpaolo Busan,^c  Nicola Molinaro,^{d,e}  and Mikel Lizarazu^d 

^aDepartment of Developmental and Social Psychology, University of Padua, Italy ^bPadova Neuroscience Center, University of Padua, Italy
^cDepartment of Medical, Surgical and Health Sciences, University of Trieste, Italy ^dBasque Center on Cognition, Brain and Language, Donostia-San Sebastián, Spain ^eIkerbasque, Basque Foundation for Science, Bilbao, Spain

ARTICLE INFO

Article History:
Received April 4, 2024
Revision received June 19, 2024
Accepted August 15, 2024

Editor-in-Chief: Cara E. Stepp
Editor: Soo-Eun Chang

https://doi.org/10.1044/2024_JSLHR-24-00227

ABSTRACT

Purpose: The purpose of this study was to investigate cortical tracking of speech (CTS) in adults who stutter (AWS) compared to typically fluent adults (TFAs) to test the involvement of the speech-motor network in tracking rhythmic speech information.

Method: Participants' electroencephalogram was recorded while they simply listened to sentences (listening only) or completed them by naming a picture (listening for speaking), thus manipulating the upcoming involvement of speech production. We analyzed speech-brain coherence and brain connectivity during listening.

Results: During the listening-for-speaking task, AWS exhibited reduced CTS in the 3- to 5-Hz range (theta), corresponding to the syllabic rhythm. The effect was localized in the left inferior parietal and right pre/supplementary motor regions. Connectivity analyses revealed that TFAs had stronger information transfer in the theta range in both tasks in fronto-temporo-parietal regions. When considering the whole sample of participants, increased connectivity from the right superior temporal cortex to the left sensorimotor cortex was correlated with faster naming times in the listening-for-speaking task.

Conclusions: Atypical speech-motor functioning in stuttering impacts speech perception, especially in situations requiring articulatory alertness. The involvement of frontal and (pre)motor regions in CTS in TFAs is highlighted. Further investigation is needed into speech perception in individuals with speech-motor deficits, especially when smooth transitioning between listening and speaking is required, such as in real-life conversational settings.

Supplemental Material: <https://doi.org/10.23641/asha.27234885>

Developmental stuttering (DS; also known as childhood-onset fluency disorder; American Psychiatric Association, 2013) is a neurodevelopmental disorder affecting the normal flow of speech and is 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 in adulthood, impairing the quality of life of affected individuals (Craig et al., 2009; Nang et al., 2018). DS likely has a multifactorial origin,

comprising motor, linguistic, emotional, neural, and genetic factors (Smith & Weber, 2017). In particular, in recent years, genetic factors have been identified (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), which may facilitate the appearance of atypical structure, function, and connectivity patterns of the central nervous system (Alm, 2021a; Craig-McQuaide et al., 2014; Etchell et al., 2018), also expressed as altered sensorimotor (SM) brain rhythms (Etchell et al., 2016; Ghaderi et al., 2018; Jenson et al., 2018, 2020; Joos et al., 2014; Saltuklaroglu et al., 2017). Such atypical neural traits lead to impaired capacities in SM planning and execution of speech (Alm, 2021b; Chang et al., 2019). Deficits in internal timing and motor

Correspondence to Simone Gastaldon: simone.gastaldon@unipd.it.
Disclosure: The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.

coordination seem to be at the core of DS (Alm, 2004), affecting wide neural systems and comprising regions such as the basal ganglia, the supplementary motor area (SMA), the inferior frontal cortex, and temporal regions (Busan, 2020; Busan et al., 2019; Craig-McQuaide et al., 2014; Etchell et al., 2018; Watkins et al., 2008). Within this network, disrupted auditory–motor interactions have been reported for DS, with difficulties in integrating auditory feedback during speaking potentially contributing to overt stuttering behaviors (Bradshaw et al., 2021; Chang et al., 2016; Daliri & Max, 2015, 2018; Halag-Milo et al., 2016; Hesse, 2023; Kim et al., 2020). Interestingly, disrupted SM function in DS has also been associated with impairments in aspects of speech perception, specifically weaker or less efficient predictive processing during spoken language comprehension (Gastaldon et al., 2023). To evaluate the role of inefficient SM function in DS during speech listening, this study considers cortical tracking of speech (CTS).

CTS, sometimes also referred to as “speech–brain entrainment” (Obleser & Kayser, 2019), is the temporal alignment of internal low-frequency brain rhythms (delta [0.1–3 Hz] and theta [4–7 Hz] frequency bands) to acoustic energy fluctuations (envelope) of the external speech signal. CTS is considered to be a valuable index reflecting the efficiency of neural processing of quasirhythmic components of speech, especially of prosodic (delta) and syllabic (theta) information (Assaneo & Poeppel, 2018; Molinaro & Lizarazu, 2018; Poeppel & Assaneo, 2020; Poeppel & Teng, 2020); disrupting such alignment can interfere with speech intelligibility (Riecke et al., 2018). Importantly, a growing body of evidence supports the view that frontal, motor, and premotor regions modulate CTS in the auditory cortex in a top-down manner (Keitel et al., 2018; Park et al., 2015). Evidence also suggests that there is a preferred frequency range at which activity in the motor and auditory cortices are coupled during speech listening. Such range lies within the theta band, specifically between 3 and 5 Hz, with a peak at 4.5 Hz (Assaneo & Poeppel, 2018). This range has been associated with the rate of both production and perception of syllabic rhythm across languages (Ding et al., 2017; Poeppel & Assaneo, 2020). In this scenario, it has been proposed that the motor system generates temporal predictions via efferent motor signals, leading to phase resetting in auditory cortices and hence to the optimization of auditory perception (Rimmele et al., 2018). Recent behavioral evidence supports this account: Higher individual speech production rates (higher fluency) and stronger auditory–motor synchronization (alignment of speech production to an external rhythm) were associated with better performance in a speech comprehension task (Lubinus et al., 2023).

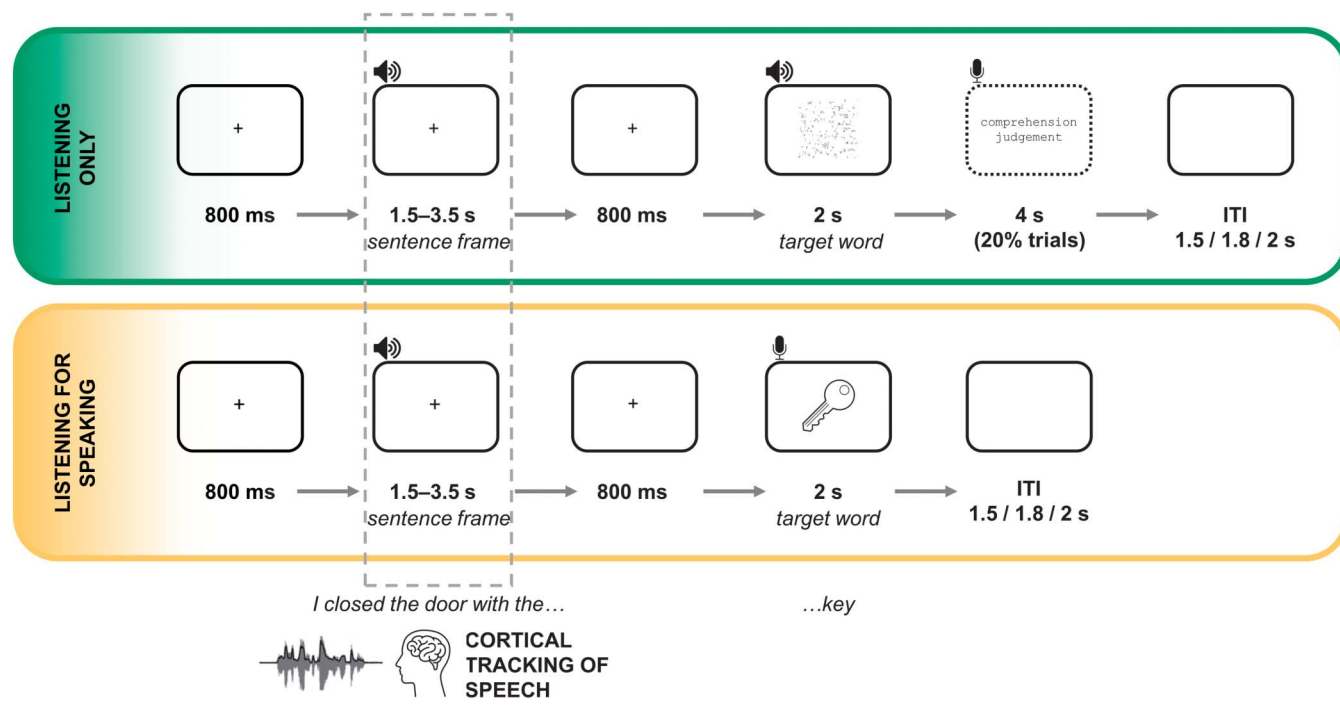
Importantly, altered brain processes related to CTS have been proposed as a risk factor for the appearance of developmental speech and/or language disorders

(Di Liberto et al., 2018; Lizarazu et al., 2015; Lizarazu, Lallier, et al., 2021; Molinaro et al., 2016; Nora et al., 2024). It has been proposed that such deficits are linked to abnormalities in the cortico–basal–thalamo–cortical circuitry involved in the processing of sensory rhythmic cues (such as beats in music and/or linguistic meter in speech), thus playing a role in processing and predicting events in a sequence (Ladányi et al., 2020). This may also be the case 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 improving our comprehension of (a) neuropathological mechanisms related to stuttering (both in the contexts of speech production and perception) and (b) the neural mechanisms involved in typical speech perception and production (and in their possible mutual interactions).

To address both these issues, 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 (TFAs; no diagnosis of speech disorders). Specifically, we measured speech–brain coherence on electroencephalogram (EEG) data, at both 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 (passively listening to the entire sentence; listening-only), in order to assess whether the upcoming involvement of the speech-motor network may have modulatory effects on CTS (see Figure 1 and the Materials and Method section). While not directly simulating everyday dyadic 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 the demands of conversational and turn-taking settings. We expected to find group differences in a restricted range within the theta band, at which activity in the auditory and motor regions is supposed to be inherently coupled and to contribute to the tracking of syllabic rhythm.

Further insight can be obtained by studying how cortical regions interact with each other during speech listening. Standard neural models of speech processing describe the presence of a dorsal and a ventral cortical stream, jointly with a cortico–subcortical loop involving the cerebellum, the thalamus, and the basal ganglia (Friederici, 2012; Hickok & Poeppel, 2004, 2007; Kotz & Schwartze, 2010; Scott & Johnsrude, 2003). While the ventral stream (comprising the auditory cortex, the anterior and posterior temporal lobe, and the inferior frontal cortex) is primarily tasked with mapping auditory information onto

Figure 1. Experimental design. Participants listened to sentence frames and target final words in the listening-only task (with occasional true/false judgment questions), 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 at picture onset in the listening-for-speaking task were collected.



linguistic representations during speech comprehension, the dorsal stream (comprising the auditory cortex; inferior parietal regions; the posterior temporal cortex; motor, premotor, and supplementary motor regions; and the inferior frontal cortex) is concerned with mapping auditory and motor representations of linguistic sounds, primarily during speech production. However, the dorsal stream seems to be useful also during speech listening, especially in challenging listening situations (Perron et al., 2024; Skipper et al., 2017), thanks to the role of this network in providing temporal structure for processing, jointly with the basal ganglia (Kotz & Schwartz, 2010). This should be particularly true when specifically targeting the tracking of rhythmic information in the speech stream, which requires sensibility to temporal regularities (Ladányi et al., 2020), as is the case of the present study. For these reasons, we also analyzed brain connectivity by using partial directed coherence (PDC) to quantify directional neural interactions between brain regions implicated in SM and speech processing. This analysis allows us to investigate frequency-specific directional communication between cortical regions during envelope tracking. Previous studies that looked at frequency-specific brain connectivity during speech envelope tracking in various contexts highlighted a variety of connectivity patterns. Importantly, many of such patterns involved regions part of the dorsal stream in the delta and theta

frequency ranges (e.g., Becker & Hervais-Adelman, 2023; Giordano et al., 2017; Hincapié Casas et al., 2021; Lizarazu, Lallier, et al., 2021; Molinaro et al., 2016; Park et al., 2015), supporting the idea that such network is relevant for tracking rhythmic patterns in the speech stream.

Based on the models and previous findings on frequency-specific connectivity in envelope tracking, we hypothesize that neural connectivity is reduced in stuttering especially in the dorsal stream, suggested to be responsible for auditory–motor transformations. 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., 2018, 2022; Sommer et al., 2002; Watkins et al., 2008). In fact, DS is characterized by abnormal structural and functional connectivity of large neural circuits, especially involving SM and speech–brain regions (for a recent review, see Etchell et al., 2018). For example, concerning measures of white matter integrity and efficiency, Sommer et al. (2002) showed that DS is characterized by lower fractional anisotropy in fibers below cortical regions representing larynx and tongue muscles in the left primary motor and somatosensory cortices, suggesting disconnection among speech/motor cortical areas. Watkins et al. (2008) found white matter deficits

in regions underlying premotor cortices, suggesting impairments in connectivity with temporal regions, primary motor cortex, and inferior parietal cortex, thus weakening integration of speech and sensory feedback, as well as execution of articulatory movements in stuttering. Neef et al. (2018) disentangled the (adaptive or maladaptive) compensatory role of right-hemisphere speech-motor regions in DS showing that stuttering severity correlates with connectivity of right frontal (hyperactive) regions, as well as with fractional anisotropy of the left superior longitudinal fascicle after speech therapy (Neef et al., 2022). When considering functional evidence (especially obtained from EEG data, as employed in this work), Busan et al. (2019) suggest that DS is characterized by a delay in recruiting left+hemisphere frontoparietal networks, followed by activation in homologous regions of the right hemisphere, thus suggesting a possible neural substrate for the emergence of dysfluencies. Along this line, Caruso et al. (2023) found reduced connectivity of the left SM circuits in motor tasks in stuttering. More specifically, interhemispheric connectivity was weaker at lower frequencies (delta and theta ranges) and stronger in the beta band, suggesting that the right hemisphere might be recruited to support SM processing in DS and that an altered balance of the neural activity might be a fundamental aspect of stuttering.

All this considered, a better understanding of neural connectivity is fundamental also in the context of the present work, especially when hypothesizing the relevance of auditory–motor pathways in tracking rhythmic information during speech listening. However, we remain agnostic as to specific patterns (directionality) and potential additional group differences (e.g., stronger connectivity in AWS in other pathways, reflecting compensatory mechanisms). To this extent, the PDC analysis is partially theoretically driven (identify pathways compatible with a dorsal processing stream) and partially exploratory. In summary, given the picture outlined above, we can hypothesize the following.

1. CTS may be reduced in AWS relative to TFAs, 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 (see the Sensor-Level Analysis section).
2. In AWS, reduced CTS may be found in speech-motor and premotor regions, in addition to auditory and associative regions (see the Source-Level Analysis section).
3. In AWS, regions that are considered to be part of the dorsal stream (inferior frontal cortex, premotor

and supplementary motor regions, SM and temporo-parietal regions) may be communicating less efficiently with auditory regions during speech tracking, thus displaying reduced connectivity (see the PDC Analysis section).

Materials and Method

Participants

We analyzed CTS in a data set collected for a previous study on spoken sentence processing in AWS, which focused on different time windows, used different analyses, and had different aims (Gastaldon et al., 2023). The original study included 14 right-handed male AWS and 14 right-handed typically fluent male adults (TFAs). The participants were matched for age and handedness. All participants were native speakers of Italian. The original study was approved by the Ethical Committee for Psychological Research of the University of Padova (Protocol No. 3073) 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, four participants were excluded due to excessively noisy EEG data during sentence frame presentation for the analyses conducted here. The remaining 24 participants, 12 AWS and 12 TFAs, were matched for age (AWS: $M_{\text{age}} = 34.44$ years, $SD = 9.37$; TFAs: $M_{\text{age}} = 33.42$ years, $SD = 8.94$) and handedness (AWS: $M = 83.75$, $SD = 20.57$; TFAs: $M = 85.00$, $SD = 23.06$), as assessed by means of the Edinburgh Handedness Inventory (Oldfield, 1971). From the original study, we also retrieved data about the Stuttering Severity Index–Fourth Edition (SSI-4; Riley, 2009) of each adult who stutters, which assigns an overall individual score by considering various characteristics of stuttering events and the occurrence of concomitant movements associated with disfluencies (e.g., oro-facial and limb movements) in both reading and spontaneous speech tasks. The higher the SSI-4 score, the more severe the stuttering is. The final set was composed of six participants with very mild severity, four with mild, and two with severe. Exploratory correlations between SSI-4 and CTS and connectivity data were performed (statistical threshold for explorative correlations: $p \leq .01$, two-tailed).

Stimuli and Procedure

The stimuli were the same with that used in the studies of Gastaldon et al. (2020, 2023; see the Open Science Framework [OSF] repository for additional information on stimuli characteristics: <https://osf.io/tcbsh/>). They consisted of 256 sentence frames (sentences without the final word,

ranging in duration from 1.55 to 3.54 s; mean duration = 2.39, $SD = 0.4$), which were paired with 128 target words and 128 black and white line pictures (124×124 pixels), such that each word and picture appeared twice, completing a high and low constraint sentence frame. Spoken stimuli were uttered by a female native Italian speaker, recorded and digitized at 44.1 kHz using Audacity. Audio files (*.wav) were also segmented using Audacity. During the task, participants listened to the sentence frames; then after an 800-ms pause, they either heard a word (listening-only task) or had to produce it by naming a picture (listening-for-speaking task), in two distinct blocks, which were counterbalanced across participants. In the listening-only task, true/false comprehension questions were asked at the end of the trial in 20% of the trials, to maintain the participant's engagement (see also Figure 1). Due to the aims of the original study (Gastaldon et al., 2023), half of the sentence frames induced high constraining contexts for the final word and half induced low constraining contexts, thus manipulating target word predictability. However, in order to allow for a better estimation of CTS and increase signal-to-noise ratio (SNR) and statistical power (see below), we did not divide the sentence frames into high versus low constraining contexts in the present analyses. We recognize that this may be a highly relevant variable that should be investigated in future studies (for a study in the normal population, see Molinaro et al., 2021); however, here we were limited in terms of SNR and number of trials. Note that groups did not differ in the number of EEG data segments coming from high and low constraining sentences in either task, thus excluding the possibility of any effect of this factor on any group differences (see Supplemental Material S5). Therefore, here we focused on the manipulation of task demands, which implied two different listening conditions: listening for comprehension (listening-only) or listening in order to complete the sentence as quickly as possible by naming a picture (listening-for-speaking). Participants sat in a dimly lit room in front of a computer screen. The experimental material was delivered through E-Prime 2.0 (Psychology Software Tools). Auditory stimuli were presented through built-in speakers. Responses (picture naming and true/false answers) 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 s. 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 EEG was recorded using a BrainAmp amplifier and BrainVision Recorder software (BrainProducts). 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 bandpass filtered online from 0.1 to 1000 Hz.

The preprocessing pipeline for the present work 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 one to four and one to three 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. As noted above, this led to the exclusion of two participants from each group, resulting in a final sample of 24 participants (12 AWS and 12 TFAs). In the case of TFAs, 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 $T_s < 1.6$, all $p_s > .11$, two-tailed t test). EEG data and MATLAB scripts for the analyses described in the following sections are available on a dedicated OSF repository: <https://osf.io/7gpyb/>.

CTS Analysis

Sensor-Level Analysis

Coherence measures the degree of phase synchronization between two signals in the frequency domain. For each participant and condition, we used coherence to quantify the CTS, which represents the coupling between the speech temporal envelope and cortical oscillations. We obtained the envelope of the speech signal from the Hilbert transformed broadband stimulus waveform. According to previous research in speech processing, we expected to find strong CTS in the low-frequency (< 10 Hz) spectrum and in temporal sensors (Ershaid et al., 2024; Issa et al., 2024; Lizarazu, Scotto di Covella, 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, six channels allocated over the left hemisphere (C3, C5, CP3, CP5, FC3, FC5) and additional six over the right hemisphere (C4, C6, CP4, CP6, FC4, FC6). Artifact-free trials were segmented into 1-s windows with 50% overlap. Coherence was then calculated using the cross-spectral density (CSD) of the FFT of the two signals (i.e., speech envelope and EEG data segments), normalized by the power spectrum of each signal. For each EEG sensor, coherence was calculated in the 1- to 15-Hz frequency band with a 1-Hz (inverse of the segment duration) frequency resolution (Molinaro et al., 2016; Molinaro & Lizarazu, 2018). This procedure was followed for each participant and task/listening condition.

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 data set and with an equal number of trials as the actual EEG–audio pairing data set.

For the statistical analysis, we calculated the mean CTS values (z -scored coherence) within the theta band, specifically in the 3- to 5-Hz frequency range. We focused on this frequency range because of two specific reasons: (a) A peak is present in our auditory stimuli in the same frequency range, indicating syllabic rhythm (see Supplemental Material S2), and (b) the existence of a frequency-restricted preference for the coupling between auditory and motor regions, as explained in the first part (see also Assaneo & Poeppel, 2018). To assess group differences in each task, we conducted an analysis of variance (ANOVA) on the z -transformed coherence values, with hemisphere (left vs. right) as the within-subject factor and group (TFAs vs. AWS) as the between-subjects factor (considering effects of main factors and their interaction; post hoc analyses conducted using t test; statistical threshold at $p \leq .05$, two-tailed).

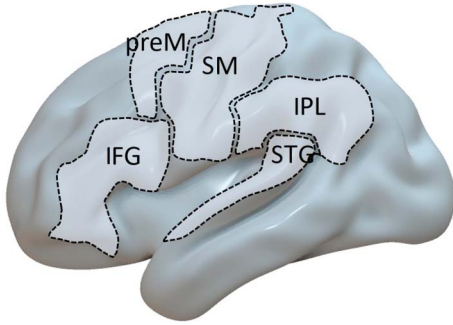
Source-Level Analysis

Coherence values were also estimated at the source level for each participant and condition in the theta band (3–5 Hz), where significant results were observed at the scalp level. For the source-level analysis, 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 three-dimensional surfaces (skin, skull, brain) derived from the MNI data set. 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 CSD matrix. The CSD matrices were computed within the theta (4 Hz with ± 1 Hz frequency smoothing) band by applying the fast Fourier transform to 1-s data segments in sliding windows shifting in 0.5-s 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, we selected five regions of interest (ROIs) from the Automatic Anatomical Labeling (Tzourio-Mazoyer et al., 2002). More specifically, ROIs were defined considering that stuttering mainly affects neural networks that are fundamental for SM processing, thus impairing speech planning, programming, and execution (compare with Chang et al., 2019). In this context, abnormal neural activity in areas such as the inferior frontal cortex, primary somato-motor regions, auditory cortex, SMA, premotor cortex, and associative regions (such as the parietal cortex) have been consistently reported as neural markers of DS (see Belyk et al., 2015, 2017; Brown et al., 2005; Budde et al., 2014; Busan, 2020; Busan et al., 2019; Chang & Guenther, 2020; Chang et al., 2019; Craig-McQuaide et al., 2014; Etchell et al., 2018; Ingham et al., 2012; Neef et al., 2015; Zhang et al., 2022). Therefore, within each cerebral hemisphere, we defined the subsequent “clusters” of brain ROIs (as shown in Figure 2):

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 hemispheres: (a) the inferior frontal gyrus (IFG, comprising the pars opercularis, triangularis, and orbitalis), (b) the premotor and supplementary motor cortex (preM), (c) the sensorimotor (SM) strip, (d) the inferior parietal lobule (IPL), and (e) the superior temporal gyrus (STG).



(a) the inferior frontal gyrus (IFG; comprising the pars opercularis, pars triangularis, and pars orbitalis), (b) the premotor and supplementary motor cortex (preM), (c) the SM strip (comprising the precentral and postcentral gyri), (d) the inferior parietal lobule (IPL; comprising the supramarginal and angular gyri), and (e) the superior temporal gyrus (STG). For each task, we employed the Wilcoxon ranked sum nonparametric test to assess group differences on the mean of the z -scored coherence values within each ROI (statistical threshold at $p \leq .05$, two-tailed).

PDC Analysis

We employed 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. 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,

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

$$\begin{pmatrix} s_1(t) \\ \vdots \\ s_M(t) \end{pmatrix} = \sum_{r=1}^p a_r \begin{pmatrix} s_1(k-r) \\ \vdots \\ s_M(k-r) \end{pmatrix} + \begin{pmatrix} \varepsilon_1(t) \\ \vdots \\ \varepsilon_M(t) \end{pmatrix}, \quad (2)$$

where $s(t) = (s_1(t), s_2(t), \dots, s_M(t))$ are the stationary M -dimensional simultaneously measured time series in each ROI, a_r are the $M \times M$ coefficient matrices of the model, and $\varepsilon(t)$ 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.

$$A(f) = I - \sum_{r=1}^p a(r)e^{-i2\pi fr/p} \quad (3)$$

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

$$\text{PDC}_{j \rightarrow i}(f) = \frac{|A_{ij}(f)|}{\sqrt{\sum_k |A_{kj}(f)|^2}}. \quad (4)$$

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), and the model order was computed using algorithms developed in the Multivariate Autoregressive Model Fitting (ARfit) software package (Schneider & Neumaier, 2001). To assess group differences, separately for each task, we used the Wilcoxon ranked sum nonparametric test on PDC values (statistical threshold at $p \leq .05$, two-tailed).

Naming Accuracy and Response Times Analysis (Listening-for-Speaking Task)

For naming latencies (response times [RTs]), 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-s *.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-s recording (in order to be able to assess the correctness of the response).

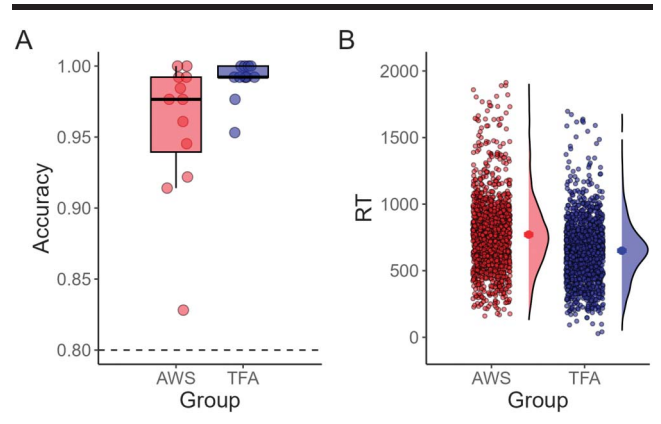
Statistical analyses were performed in R. Accuracy was analyzed with a generalized linear mixed-effects model (GLMM) with binomial distribution family. Group, lexical frequency of the target word (retrieved through PhonItalia; Goslin et al., 2014), and repetition (the same target picture was presented twice in the task, associated with two different sentence frames) were set as fixed effects, while participant and item were set 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 were set as random intercepts. As explained above, we decided not to include sentence constraint as a factor here since, for the coherence analysis (the main focus of the present work), we did not differentiate between the two conditions for methodological reasons. GLMM was 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 $p \leq .01$, two-tailed).

Results

Naming (Listening-for-Speaking Task)

Accuracy and RTs are shown in Figure 3, whereas model summaries are reported in Table 1. Participants of both groups had a very high accuracy in producing the

Figure 3. Behavioral results for the listening-for-speaking task. (A) Subject-level (individual) accuracy scores (dots) and box plots; y-axis starting at 0.8. (B) Single-trial response times (dots), group-level means with error bars, and density distributions. AWS = adults who stutter; RT = response time; TFAs = typical fluent adults.



correct word (AWS: $M = 0.96$, $SD = 0.2$; TFAs: $M = 0.99$, $SD = 0.1$; see Figure 3A). However, the model revealed a main effect of repetition (higher accuracy when the picture appeared for the second time) and a main effect of group, with AWS less accurate than TFAs (see Table 1). Regarding RTs, AWS were slower than TFAs (AWS: $M = 771.19$ ms, $SD = 267.77$; TFAs: $M = 650.53$ ms, $SD = 219.98$; see Figure 3B). The model revealed a main effect of repetition and, importantly, a main effect of group (see Table 1). To test the robustness of the results to possible outliers for accuracy, we rerun the analysis by excluding the AWS participant with accuracy = 0.84 (see Figure 3), and the results are still consistent (main effect of group: $t = -2.35$, $p = .019$, 95% CI [0.33, 0.91]). Speculatively, lower accuracy, rather than reflecting possible inefficiency in retrieving lexical items in AWS, is likely due to the limited time available for recording the responses (2 s after picture onset): Sometimes, AWS may have provided the correct response outside this window, making it impossible to

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

Predictor	Accuracy			
	Estimates	CI	Statistics	p
(Intercept)	60.83	[20.21, 183.11]	7.31	< .001
Lexical frequency	1.22	[0.97, 1.54]	1.68	.093
Group	0.47	[0.27, 0.82]	-2.64	.008
Repetition	0.70	[0.55, 0.90]	-2.84	.005
Predictor	Response times			
	Estimates	CI	Statistics	p
(Intercept)	730.15	[688.82, 771.48]	34.64	< .001
Lexical frequency	-6.05	[-14.21, 2.10]	-1.45	.146
Group	60.43	[34.44, 86.41]	4.56	< .001
Repetition	36.44	[29.12, 43.76]	9.76	< .001

Note. CI = confidence interval.

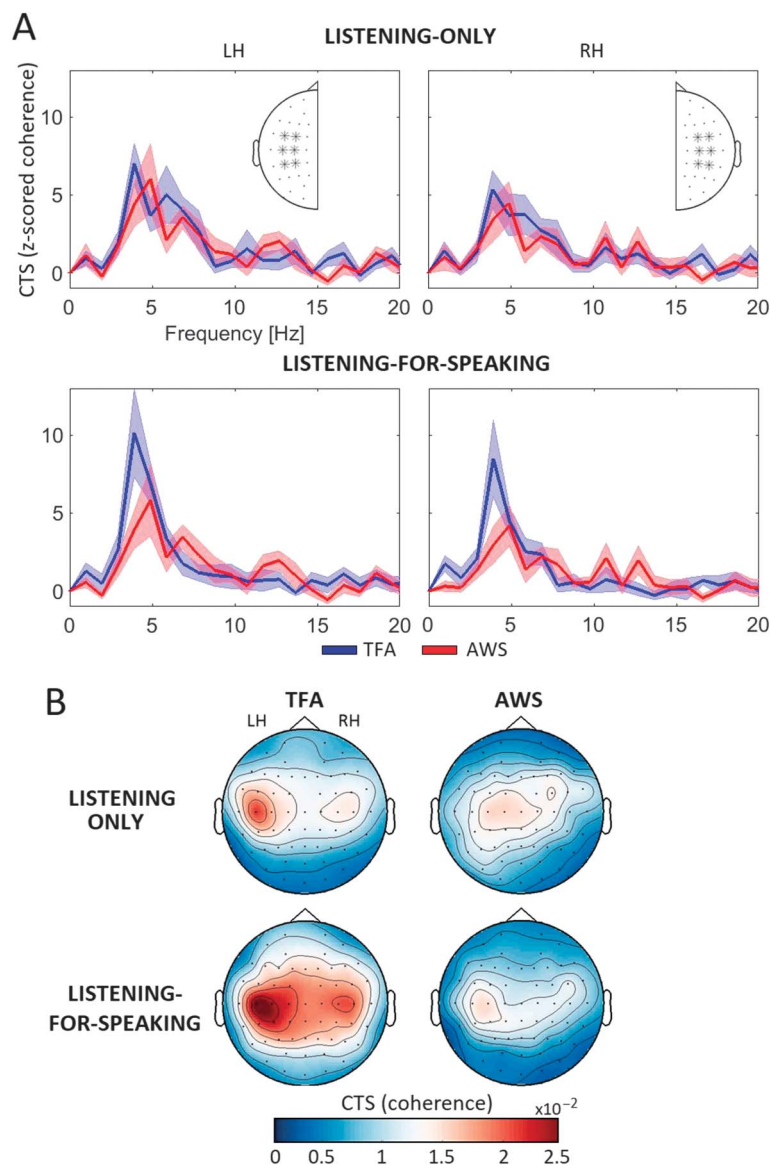
evaluate their response off-line, hence the reduced accuracy (i.e., fewer trials coded as correct). This interpretation is compatible with the fact that the accuracy outlier in the AWS group (accuracy = 0.84) is also the one with longest mean RTs for correct responses (RT = 1,155 ms).

Sensor-Level CTS

Initially, we conducted an assessment of sensor-level CTS within the 1- to 15-Hz frequency range for

each group (TFAs and AWS) and task (listening-only and listening-for-speaking). Consistent with previous studies, we observed that during speech listening, CTS was highest in the theta (3–5 Hz) frequency band (see Figure 4A) in bilateral frontocentral, central, and centro parietal sensors (see Figure 4B), consistent with the topography usually found in magnetoencephalography and electroencephalography studies on coherence as measure of CTS in the theta range (Destoky et al., 2019).

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- to 15-Hz frequency range can be observed across representative sensors (C3, C5, CP3, CP5, FC3, FC5, C4, C6, CP4, CP6, FC4, and FC6) of the LH and the RH. (B) For each group (TFA and AWS) and task (listening-only and listening-for-speaking), we plotted the topographic maps of uncorrected coherence values in the theta (3–5 Hz) frequency band. AWS = adults who stutter; CTS = cortical tracking of speech; EEG = electroencephalogram; LH = left hemisphere; RH = right hemisphere; RT = response time; TFA = typical fluent adult.



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 $F_s(1, 22) < 1.97$, all $p_s > .17$, $\eta^2_s < .06$). However, we did observe a main effect of group, $F(1, 22) = 4.07$, $p = .05$, $\eta^2 = .15$, in the CTS values for the listening-for-speaking task. Post hoc tests showed that CTS was significantly higher in TFAs compared to AWS ($t = 2.02$, $p = .05$, Cohen's $d = 0.80$). No statistically significant correlations with RTs or SSI-4 were found.

Source-Level CTS

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

In line with the results observed at the sensor level, we found that the CTS values were significantly stronger for individuals with TFAs compared to AWS in the left IPL ($M_{TFA} = 1.92$, $SD_{TFA} = 2.58$; $M_{AWS} = 0.22$, $SD_{AWS} = 0.34$; $p = .03$) and in the right preM regions ($M_{TFA} = 0.07$, $SD_{TFA} = 1.21$; $M_{AWS} = -0.74$, $SD_{AWS} = 0.36$; $p = .01$), only for the listening-for-speaking task (see Figure 6). We did not find any group differences in the listening-only task (all $p > .09$; see Supplemental Material

S1). No statistically significant correlations with RTs and SSI-4 were found.

Source-Level Connectivity

We used PDC to assess causal functional connectivity during speech processing in the theta band (3–5 Hz) between different ROIs. For the listening-only task (see Figure 7), we observed significantly higher connectivity in TFAs compared to AWS from the left STG to the right IFG ($M_{TFA} = 0.12$, $SD_{TFA} = 0.05$; $M_{AWS} = 0.07$, $SD_{AWS} = 0.04$; $p = .01$) and from the right IFG to the left IPL ($M_{TFA} = 0.03$, $SD_{TFA} = 0.01$; $M_{AWS} = 0.02$, $SD_{AWS} < 0.01$; $p = .04$). For the listening-for-speaking task (see Figure 5), we observed significantly higher connectivity in TFAs compared to AWS from the right STG to the left IPL ($M_{TFA} = 0.05$, $SD_{TFA} = 0.06$; $M_{AWS} = 0.03$, $SD_{AWS} = 0.02$; $p = .05$) and from the right STG to the left SM regions ($M_{TFA} = 0.11$, $SD_{TFA} = 0.06$; $M_{AWS} = 0.06$, $SD_{AWS} = 0.04$; $p = .03$). Interestingly, when considering both groups together, 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 ($r = -.56$, $p = .0048$): Stronger directional connectivity between these regions is associated with faster RTs (see Figure 8).

Discussion

In the present work, we analyzed CTS in a group of AWS and a group of TFAs. To investigate the role of the alertness state of the speech-motor system in CTS, we

Figure 5. Source reconstruction of the CTS values in the theta range. For each group (TFA and AWS) and listening condition (listening-only and listening-for-speaking), we plotted the source maps of CTS values (coherence values converted into z score using the mean and standard deviation derived from the CTS values in all the sources) in the theta (3–5 Hz) frequency band. AWS = adults who stutter; CTS = cortical tracking of speech; LH = left hemisphere; RH = right hemisphere; TFA = typical fluent adult.

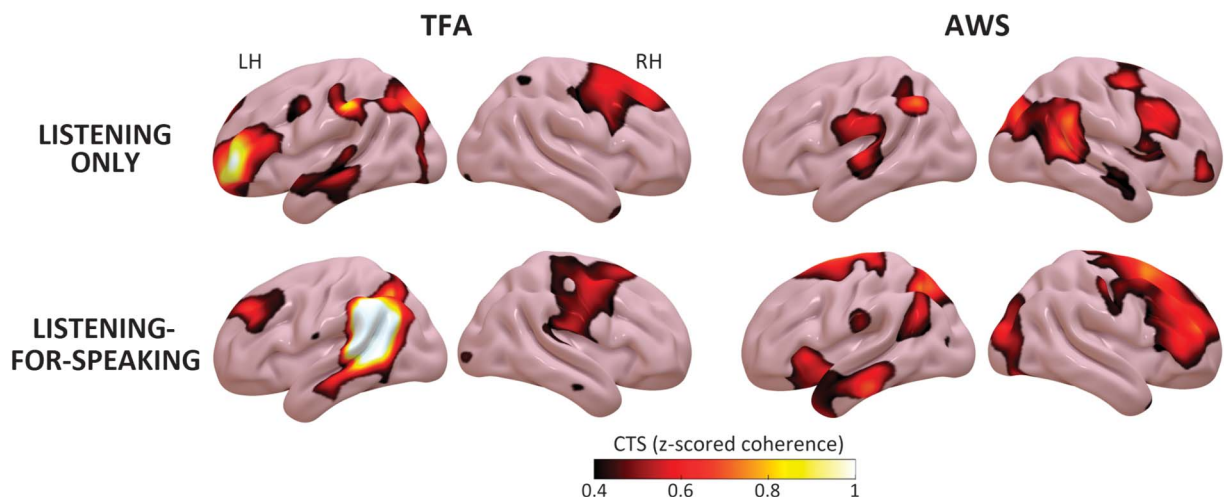
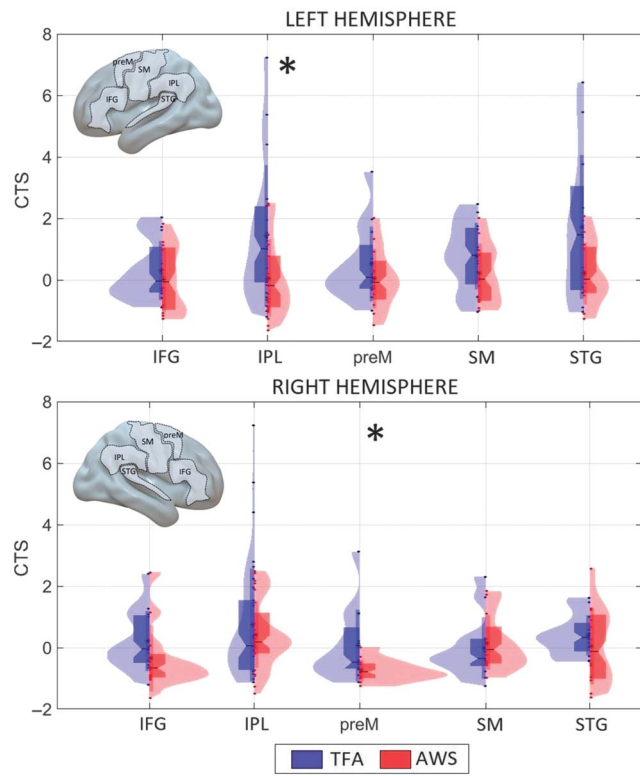
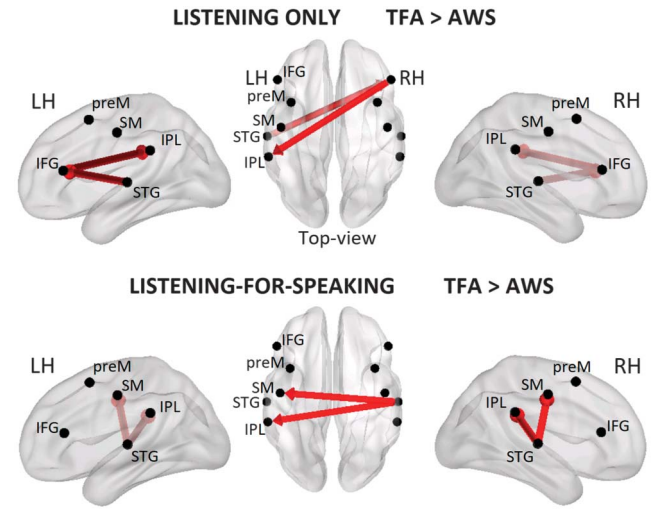


Figure 6. Mean CTS in the 3–5 Hz range (theta band) in each region of interest for each group. Box plots are overlaid with individual data points and density distributions. Each dot represents data from one of the participants (blue for TFA and red for AWS). Boxes cover the 25th to the 75th percentile (IQR). The middle of the box represents the median. Whiskers extend from the 25th to the 75th percentile to cover all data points lying within 1.5 times the IQR (from the 25th and the 75th percentile, respectively). Regions showing a significant group effect are marked with an asterisk. AWS = adults who stutter; CTS = cortical tracking of speech; IFG = inferior frontal gyrus; IPL = inferior parietal lobule; IQR = interquartile range; preM = premotor and supplementary motor cortex; RH = right hemisphere; SM = sensorimotor; STG = superior temporal gyrus; TFA = typical fluent adult.



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 TFAs in the listening-for-speaking task at both the sensor (bilaterally around the temporal regions) and the neural source levels. More specifically, at source level, AWS showed lower CTS in the left inferior parietal/temporoparietal cortex and in the right premotor and supplementary motor regions. Cortical connectivity measures in the theta range were differently modulated for AWS as well, with weaker connections in both listening conditions, indicating lower inter-hemispheric information exchange between frontal, auditory/temporal, and SM regions. Notably, in the listening-for-

Figure 7. Causal functional connectivity analysis. For each task, we included the connections that exhibited statistically significantly higher partial directed coherence values for TFA compared to AWS. We included a seed for each of the regions of interest in both the LH and the RH. AWS = adults who stutter; IFG = inferior frontal gyrus; IPL = inferior parietal lobule; LH = left hemisphere; preM = premotor and supplementary motor cortex; RH = right hemisphere; SM = sensorimotor; STG = superior temporal gyrus; TFA = typical fluent adult.

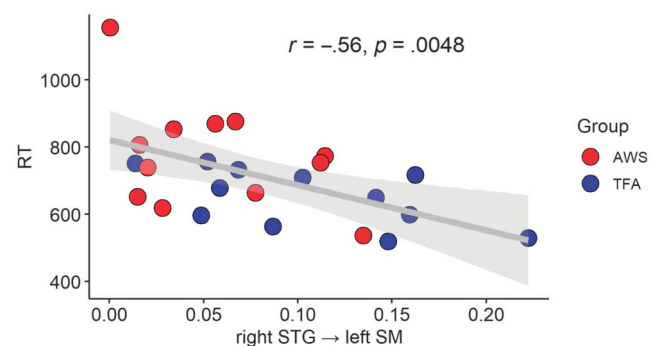


speaking task, we also found slower (speech) RTs in AWS and a significant negative correlation between RTs and connectivity from the right STG to the left SM cortex when considering all participants, reinforcing the arguments we lay out next.

Cortical Tracking of Syllabic Rhythm Is Reduced in DS When Listening for Speaking

The listening-for-speaking condition in this study required speech listening to be interwoven with speech

Figure 8. Connectivity–RT correlation (listening-for-speaking). Scatterplot showing the correlation between RTs and connectivity from the right superior temporal gyrus to the left sensorimotor cortex. AWS = adults who stutter; RT = response time; SM = sensorimotor; STG = superior temporal gyrus; TFA = typical fluent adult.



production, similar to turn-taking in conversational settings (Levinson, 2016). This entails the ability to efficiently time the transition between listening and speaking and appropriately plan production initiation while still attending to speech. The present findings suggest that CTS in AWS is impaired especially in such situations. As highlighted in the first part, CTS is a neural index reflecting 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). Crucially, coupling between oscillatory activity in auditory and speech-motor regions in a restricted frequency range within the theta band seems to support this process, specifically for the tracking of syllabic rhythm (Assaneo & Poeppel, 2018; Keitel et al., 2018; Morillon & Baillet, 2017; Park et al., 2015). In this study, we found that in a population characterized by inefficient timing and implementation of speech-motor processes, that is, AWS (Alm, 2004, 2021b; Busan, 2020; Chang & Guenther, 2020), CTS is also affected as a result of disruptions to auditory-motor coupling.

At the source level, in the listening-for-speaking task, we observed CTS reduction in the left inferior parietal cortex and in the right premotor and supplementary motor regions in AWS compared to TFAs. All these regions are key cortical substrates for speech-motor coordination. The IPL, comprising the supramarginal gyrus and the angular gyrus, has been associated with a variety of functions, including verbal working memory, auditory spatial localization, SM 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 (nonanatomically defined) temporoparietal 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 Hickok and Poeppel's model, the Sylvian parietal temporal area (located between the IPL and the posterior part of the STG, thus situated within the TPJ) is proposed to be an interface between auditory codes and motor programs supporting successful SM integration during speech production, instantiated in the dorsal pathway. The IPL is also key in neuro-computational models of speech production such as the Directions Into Velocities of Articulators (DIVA)/Gradient Order DIVA (GODIVA) models (Guenther, 2016), which propose that somatosensory error maps of the difference between intended and actual somatic states are computed in the IPL 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 (SMA and pre-SMA). In the speech-motor control literature, these regions have been highlighted in the composition and the timing of execution of speech-motor command sequences (Alario et al., 2006; Ghosh et al., 2008; Guenther, 2016). More specifically, in the DIVA/GODIVA models, the SMA is responsible for the correct initiation of stored speech-motor units, while the pre-SMA represents the global sequential structure of the syllables to be produced. On the other hand, these models propose that right-hemisphere premotor regions may be a component of a feedback/control speech-motor network (Bohland et al., 2010; Guenther, 2016; Tourville & Guenther, 2011; see Chang & Guenther, 2020; Civier et al., 2013, for a perspective on DS). Notably, rhythm processing seems to be particularly reliant upon such cortical structures (together with subcortical regions), in both the speech and nonspeech domains (Cannon & Patel, 2021; Fiveash et al., 2021; Kasdan et al., 2022; Ladányi et al., 2020). Additionally, the SMA has also been linked to the mediation of motor-sound representations in auditory prediction and speech imagery (Lima et al., 2016). Crucially, premotor and supplementary motor regions are among the regions that are found to be most dysfunctional in DS (Busan, 2020; Busan et al., 2019; Chang & Guenther, 2020; Civier et al., 2013; Etchell et al., 2018). When related with present findings, this body of evidence is compatible with a key role of premotor/supplementary motor regions in tracking rhythmic information at the syllabic level during speech perception, speculatively by transforming motor information into auditory templates for syllabic tracking.

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

The fact that we found differences in CTS within the theta range is particularly interesting from a speech-motor impairment point of view. 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; Poeppel & Assaneo, 2020). Indeed, a peak was present in the theta range in our audio stimuli (3–4 Hz), reflecting this acoustic property (see Supplemental Material S2). Importantly, 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 as to the possible configurations they can produce and the speed at which they can be executed; the syllable represents the optimal motor programming unit that the neural system can send to the motor system for execution (Guenther, 2016; Poeppel & Assaneo, 2020). Crucially, individuals with DS seems to be impaired in the ability to automatically activate syllabic motor units associated with learned sound sequences via the basal ganglia motor loop connected to the pre-SMA and SMA (Alm, 2004, 2021a, 2021b; Busan, 2020; Chang & Guenther, 2020; Civier et al., 2013).

Therefore, it appears that there is a circular relationship that, stemming from biomechanical articulation constraints via neural motor program units, leads 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 closely related to the results obtained in the present study: The speech-motor production system is involved in tracking acoustic properties that arise from articulatory-motor constraints. When such a system is unstable (as in the case of DS), perceptual tracking is also less efficient, more noticeably when listening and speaking are interwoven and partially overlapping neural resources are required, thus overburdening an already unstable system. Behaviorally, the presence of slower RTs in AWS further supports this view.

Weaker Interhemispheric Connectivity Among Bilateral Auditory and SM Regions in DS When Listening to Speech

We found weaker connectivity patterns in AWS relative to TFAs 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 SM

regions and to the left IPL/TPJ. While we do not interpret the different patterns across listening conditions, all the regions involved are consistent with a dorsal stream of processing (Friederici, 2012; Hickok & Poeppel, 2007), supporting the idea that auditory–motor mapping is important for CTS, at least of syllabic rhythm. This may be related to the nature of the syllable itself, representing the optimal motor unit for the human speech-motor system (Poeppel & Assaneo, 2020; Strauß & Schwartz, 2017). Reduced connectivity in AWS likely reflects lower availability of neural resources for information exchange between regions that are instrumental for auditory and motor processing and integration, 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 were associated with increased connectivity between these regions. This may indicate that efficiently sending rhythmic auditory information to the primary SM cortex when speech listening has to be managed with (overt) upcoming speech-motor engagement facilitates 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 TFAs, was not strongly evident in AWS ($r = -.44$, $p = .15$) but was present in TFAs ($r = -.58$, $p = .045$; see Supplemental Material S4). As a further indication, we would also like to highlight that, albeit statistically not significant, an interesting trend was present in AWS when looking at this very same connectivity pattern and SSI-4, where a negative relation is found ($r = -.55$, $p = .06$; see Supplemental Material S4 and Supplemental Material S3): Higher SSI-4 scores—hence, more severe stuttering—were associated with weaker right STG → left SM cortex connectivity.

Together, these findings strongly suggest that stuttering may be associated with weaker connectivity between auditory and SM regions, which is fundamental for cortical tracking of syllabic units, in turn leading to poorer behavioral performance in terms of RTs. This interpretation may be also compatible with proposals according to which an effective connection between these regions may be helpful for better managing (or “by-passing”) disfluencies, perhaps by exploiting rhythmic or tracked cues (Etchell et al., 2014). Consistent with this suggestive although marginal evidence, activity in right-hemispheric frontotemporal regions is often reported as neural markers of stuttering trait and state (e.g., Belyk et al., 2015, 2017; Brown et al., 2005; Budde et al., 2014; Craig-McQuaide

et al., 2014; Etchell et al., 2014, 2018; Ingham et al., 2012; Neef et al., 2015; Stasak et al., 2021), 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; Etchell et al., 2014; Neef et al., 2015, 2016, 2018, 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 (a) CTS requires neural resources that sustain SM processes for facilitating speech perception and intelligibility; (b) DS may lead to suboptimal CTS, especially when additional resources are needed for supporting concomitant speech preparation for upcoming production; and (c) DS not only impairs speech programming and production but is a more complex neurodevelopmental disorder. Further research should clarify the extent to which DS impacts CTS (and vice versa) and how this might affect people's everyday life, hence widening the scope of possible interventions for stuttering. This is especially important in light of recent evidence suggesting that auditory-motor coupling (and individual speech production rates) may explain performance in speech comprehension tasks (Lubinus et al., 2023). Less efficient CTS may be related to more effortful spoken language comprehension at a subtle level. This is in line with the results reported by Gastaldon et al. (2023): AWS seem less efficient at 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 (see also Gastaldon et al., 2024, on the importance of studying atypical populations for a better understanding of predictive speech processing). In conclusion, research should move toward turn-based and conversational contexts (e.g., Jackson et al., 2021; Weiss, 1995) to investigate possible subtle differences in how spoken language comprehension is achieved in the stuttering brain.

Limitations

The study provides interesting new results, suggesting future venues for CTS and DS research; however, some limitations need to be taken into consideration. First of all, sample sizes are small. The primary reason lies in the difficulty in recruiting AWS participants. This is a common problem when studying neurodevelopmental

disorders at low incidence in the population, such as DS (Jones et al., 2002). To address this, in line with increasingly relevant Open Science practices, multilab projects can be an efficient way to overcome small *N*s and to generalize or disconfirm results from individual underpowered studies and to appropriately quantify effect sizes by means of meta-analyses (Heinrich & Knight, 2020; Lange, 2020; McShane et al., 2019). Note that, by making data available, we provide material for future meta-analyses and/or reanalyses, in the spirit of Open Science.

Another limitation related to DS that should be addressed in future research is that the current study involved male participants only. Persistent DS in adulthood is highly asymmetric according to sex, with a stronger incidence in males (about 1:5 ratio; Yairi & Ambrose, 2013), making recruitment inherently unbalanced. Furthermore, sex hormones may underlie neural changes related to speech-motor control relevant for the persistence or resolution of DS in adulthood (see Neef & Chang, 2024). Thus, it would be interesting to investigate sex-related differences in neural tracking of speech in DS.

Another limitation regards localization of cortical regions. This limitation is common to all studies employing EEG. However, good estimates can still be obtained when using a sufficient number of electrodes covering all the scalp (such as 64 electrodes in the present work) by following standardized electrode placement and by imposing reliable biophysical constraints to forward and inverse solutions (Lantz et al., 2003; Michel & Brunet, 2019; Michel et al., 2004; Westner et al., 2022). Future studies may employ higher density EEG systems or magnetoencephalography, combined with individual structural scans, in order to provide a more accurate picture.

Conclusions

The present work suggests that CTS recruits (pre) motor regions and regions responsible for SM integration, as well as auditory regions, supporting views proposing an interaction between these networks in speech-language perception (Pickering & Garrod, 2013; Skipper et al., 2017), in addition to their instrumental role in orchestrating successful speech production (Guenther, 2016; Hickok et al., 2011). CTS 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.

Author Contributions

Simone Gastaldon: Conceptualization (Equal), Data curation (Supporting), Formal analysis (Equal), Investigation (Lead), Methodology (Supporting), Project administration (Lead), Visualization (Equal), Writing – original draft (Lead), Writing – review & editing (Lead). **Pierpaolo Busan:** Conceptualization (Supporting), Writing – original draft (Equal), Writing – review & editing (Equal). **Nicola Molinaro:** Conceptualization (Equal), Methodology (Supporting), Supervision (Lead), Writing – review & editing (Equal). **Mikel Lizarazu:** Conceptualization (Equal), Data curation (Lead), Formal analysis (Equal), Methodology (Lead), Project administration (Supporting), Visualization (Equal), Writing – original draft (Equal), Writing – review & editing (Equal).

Data Availability Statement

Data and scripts for the main analyses are available at the Open Science Framework repository: <https://osf.io/7gpyb/>.

Acknowledgments

Research was supported by the University of Padua, by the Basque Government through the Biodiversity & Environmental Research Center 2018–2021 program, and by the Spanish State Research Agency through Basque Center on Cognition, Brain and Language's Severo Ochoa excellence accreditation CEX2020-001010-S. Simone Gastaldon was supported by a postdoctoral research grant funded by the Fondazione CARIPARO through the PHD@UNIPD call at the University of Padova (Grant CUP_C93C23003190005). Nicola Molinaro was supported by the Spanish Ministry of Science, Innovation and University (Grants RTI2018-096311-B-I00, PDC2022-133917-I00, PCI2022-135031-2, PID2022-136991NB-I00); the Agencia Estatal de Investigación; and the Fondo Europeo de Desarrollo Regional. Mikel Lizarazu was supported by the Ramón y Cajal programme of the Spanish Ministry of Science and Universities (Grant RYC2022-035497-I). Funding sources had no role in the collection, analysis, and interpretation of data. The authors would like to thank Caroline Handley for proofreading the article.

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, 15*, Article 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*, Article 663. <https://doi.org/10.3389/fnhum.2021.661880>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5™* (5th ed.). American Psychiatric Publishing. <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), Article eaao3842. <https://doi.org/10.1126/sciadv.aao3842>
- 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>
- Becker, R., & Hervais-Adelman, A. (2023). Individual theta-band cortical entrainment to speech in quiet predicts word-in-noise comprehension. *Cerebral Cortex Communications, 4*(1), 1–13. <https://doi.org/10.1093/TEXCOM/TGAD001>
- 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*, Article 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, pp. 467–479). Elsevier. <https://doi.org/10.1016/B978-0-444-63622-5.00024-3>
- Bohland, J. W., Bullock, D., & Guenther, F. H. (2010). Neural representations and mechanisms for the performance of simple speech sequences. *Journal of Cognitive Neuroscience, 22*(7), 1504–1529. <https://doi.org/10.1162/jocn.2009.21306>
- Bradshaw, A. R., Lametti, D. R., & McGettigan, C. (2021). The role of sensory feedback in developmental stuttering: A review. *Neurobiology of Language, 2*(2), 308–334. https://doi.org/10.1162/nol_a_00036
- Brown, S., Ingham, R. J., Ingham, J. C., Laird, A. R., & Fox, P. T. (2005). Stuttered and fluent speech production: An ALE meta-analysis of functional neuroimaging studies. *Human Brain Mapping, 25*(1), 105–117. <https://doi.org/10.1002/hbm.20140>
- 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*, 64, Article 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. *Neuroscience and Biobehavioral Reviews*, 68, 319–334. <https://doi.org/10.1016/j.neubiorev.2016.02.024>
- Cannon, J. J., & Patel, A. D.** (2021). How beat perception co-opts motor neurophysiology. *Trends in Cognitive Sciences*, 25(2), 137–150. <https://doi.org/10.1016/j.tics.2020.11.002>
- Caruso, V. C., Wray, A. H., Lescht, E., & Chang, S.-E.** (2023). Neural oscillatory activity and connectivity in children who stutter during a non-speech motor task. *Journal of Neurodevelopmental Disorders*, 15(1), Article 40. <https://doi.org/10.1186/s11689-023-09507-8>
- 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*, 12, 442–450. <https://doi.org/10.1016/j.nicl.2016.08.021>
- Chang, S. E., Garnett, E. O., Etschell, 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: Current Insights. *Frontiers in Psychology*, 10, Article 3088. <https://doi.org/10.3389/fpsyg.2019.03088>
- Chow, H. M., Garnett, E. O., Li, H., Etschell, 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(3), 365–380. 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, A., Blumgart, E., & Tran, Y.** (2009). The impact of stuttering on the quality of life in adults who stutter. *Journal of Fluency Disorders*, 34(2), 61–71. <https://doi.org/10.1016/j.jfludis.2009.05.002>
- Craig-McQuaide, A., Akram, H., Zrinzo, L., & Tripoliti, E.** (2014). A review of brain circuitries involved in stuttering. *Frontiers in Human Neuroscience*, 8, Article 884. <https://doi.org/10.3389/fnhum.2014.00884>
- 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>
- Destoky, F., Philippe, M., Bertels, J., Verhasselt, M., Coquelet, N., Vander Ghinst, M., Wens, V., De Tiège, X., & Bourguignon, M.** (2019). Comparing the potential of MEG and EEG to uncover brain tracking of speech temporal envelope. *NeuroImage*, 184, 201–213. <https://doi.org/10.1016/j.neuroimage.2018.09.006>
- Di Liberto, G. M., Peter, V., Kalashnikova, M., Goswami, U., Burnham, D., & Lalor, E. C.** (2018). Atypical cortical entrainment to speech in the right hemisphere underpins phonemic deficits in dyslexia. *NeuroImage*, 175, 70–79. <https://doi.org/10.1016/j.neuroimage.2018.03.072>
- Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., & Poeppel, D.** (2017). Temporal modulations in speech and music. *Neuroscience & Biobehavioral Reviews*, 81, 181–187. <https://doi.org/10.1016/j.neubiorev.2017.02.011>
- Ershaid, H., Lizarazu, M., McLaughlin, D., Cooke, M., Simantiraki, O., Koutsogiannaki, M., & Lallier, M.** (2024). Contributions of listening effort and intelligibility to cortical tracking of speech in adverse listening conditions. *Cortex*, 172, 54–71. <https://doi.org/10.1016/j.cortex.2023.11.018>
- Etschell, 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>
- Etschell, 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, Article 467. <https://doi.org/10.3389/fnhum.2014.00467>
- Etschell, 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>
- Fiveash, A., Bedoin, N., Gordon, R. L., & Tillmann, B.** (2021). Processing rhythm in speech and music: Shared mechanisms and implications for developmental speech and language disorders. *Neuropsychology*, 35(8), 771–791. <https://doi.org/10.1037/neu0000766>
- 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., Bonfiglio, N., Vespignani, F., & Peressotti, F.** (2024). Predictive language processing: Integrating comprehension and production, and what atypical populations can tell us. *Frontiers in Psychology*, 15, Article 1369177. <https://doi.org/10.3389/fpsyg.2024.1369177>
- 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*, Article 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\(2008/07-0119\)](https://doi.org/10.1044/1092-4388(2008/07-0119))
- Giordano, B. L., Ince, R. A. A., Gross, J., Schyns, P. G., Panzeri, S., & Kayser, C. (2017). Contributions of local speech encoding and functional connectivity to audio-visual speech perception. *eLife*, *6*, Article e24763. <https://doi.org/10.7554/eLife.24763>
- 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*, *37*(3), 424–438. <https://doi.org/10.2307/1912791>
- 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>
- Guenther, F. H. (2016). *Neural control of speech*. 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*, *11*, 328–338. <https://doi.org/10.1016/j.neuroimage.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>
- Heinrich, A., & Knight, S. (2020). Reproducibility in cognitive hearing research: Theoretical considerations and their practical application in multi-lab studies. *Frontiers in Psychology*, *11*, Article 500520. <https://doi.org/10.3389/fpsyg.2020.01590>
- Hesse, T. (2023). Developmental stuttering may be caused by insufficient processing of auditory feedback. *Medical Hypotheses*, *180*, Article 111166. <https://doi.org/10.1016/J.MEHY.2023.111166>
- 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>
- Hincapié Casas, A. S., Lajnef, T., Pascarella, A., Guiraud-Vinatea, H., Laaksonen, H., Bayle, D., Jerbi, K., & Boulenger, V. (2021). Neural oscillations track natural but not artificial fast speech: Novel insights from speech-brain coupling using MEG. *NeuroImage*, *244*, Article 118577. <https://doi.org/10.1016/j.neuroimage.2021.118577>
- 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>
- Issa, M. F., Khan, I., Ruzzoli, M., Molinaro, N., & Lizarazu, M. (2024). On the speech envelope in the cortical tracking of speech. *NeuroImage*, *297*, Article 120675. <https://doi.org/10.1016/j.neuroimage.2024.120675>
- 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*, Article 105878. <https://doi.org/10.1016/j.jfludis.2021.105878>
- Jenson, D., Bowers, A. L., Hudock, R., & Saltuklaroglu, T. (2020). The application of EEG mu rhythm measures to neurophysiological research in stuttering. *Frontiers in Human Neuroscience*, *13*, Article 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*, *19*, 690–702. <https://doi.org/10.1016/J.NICL.2018.05.026>
- Jones, M., Gebski, V., Onslow, M., & Packman, A. (2002). Statistical power in stuttering research: A tutorial. *Journal of Speech, Language, and Hearing Research*, *45*(2), 243–255. [https://doi.org/10.1044/1092-4388\(2002/019\)](https://doi.org/10.1044/1092-4388(2002/019))
- Joos, K., Ridder, D. De, Boey, R. A., & Vanneste, S. (2014). Functional connectivity changes in adults with developmental stuttering: A preliminary study using quantitative electroencephalography. *Frontiers in Human Neuroscience*, *8*, Article 783. <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/NEJM0A0902630>
- Kasdan, A. V., Burgess, A. N., Pizzagalli, F., Scartozzi, A., Chern, A., Kotz, S. A., Wilson, S. M., & Gordon, R. L. (2022). Identifying a brain network for musical rhythm: A functional neuroimaging meta-analysis and systematic review. *Neuroscience & Biobehavioral Reviews*, *136*, Article 104588. <https://doi.org/10.1016/j.neubiorev.2022.104588>
- Keitel, A., Gross, J., & Kayser, C. (2018). Perceptually relevant speech tracking in auditory and motor cortex reflects distinct linguistic features. *PLOS Biology*, *16*(3), Article 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>
- Kotz, S. A., & Schwartz, M. (2010). Cortical speech processing unplugged: A timely subcortico-cortical framework. *Trends in*

- Cognitive Sciences*, 14(9), 392–399. <https://doi.org/10.1016/j.tics.2010.06.005>
- 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*, 11(5), Article e1528. <https://doi.org/10.1002/wcs.1528>
- Lange, F.** (2020). Are difficult-to-study populations too difficult to study in a reliable way? *European Psychologist*, 25(1), 41–50. <https://doi.org/10.1027/1016-9040/a000384>
- Lantz, G., Grave de Peralta, R., Spinelli, L., Seeck, M., & Michel, C. M.** (2003). Epileptic source localization with high density EEG: How many electrodes are needed? *Clinical Neurophysiology*, 114(1), 63–69. [https://doi.org/10.1016/S1388-2457\(02\)00337-1](https://doi.org/10.1016/S1388-2457(02)00337-1)
- 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. *Trends in Neurosciences*, 39(8), 527–542. <https://doi.org/10.1016/j.tins.2016.06.003>
- Lizarazu, M., Lallier, M., Bourguignon, M., Carreiras, M., & Molinaro, N.** (2021). Impaired neural response to speech edges in dyslexia. *Cortex*, 135, 207–218. <https://doi.org/10.1016/j.cortex.2020.09.033>
- 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), Article 20222410. <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, Article 598949. <https://doi.org/10.3389/fnins.2021.598949>
- McShane, B. B., Tackett, J. L., Böckenholt, U., & Gelman, A.** (2019). Large-scale replication projects in contemporary psychological research. *The American Statistician*, 73(Suppl. 1), 99–105. <https://doi.org/10.1080/00031305.2018.1505655>
- Michel, C. M., & Brunet, D.** (2019). EEG source imaging: A practical review of the analysis steps. *Frontiers in Neurology*, 10, Article 325. <https://doi.org/10.3389/fneur.2019.00325>
- Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., & Grave de Peralta, R.** (2004). EEG source imaging. *Clinical Neurophysiology*, 115(10), 2195–2222. <https://doi.org/10.1016/j.clinph.2004.06.001>
- 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., Baldin, V., Pérez-Navarro, J., Lallier, M., & Ríos-López, P.** (2021). Speech-brain phase coupling is enhanced in low contextual semantic predictability conditions. *Neuropsychologia*, 156, Article 107830. <https://doi.org/10.1016/j.neuropsychologia.2021.107830>
- 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., Büttfering, C., Schmidt-Samoa, C., Friederici, A. D., Paulus, W., & Sommer, M.** (2018). 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. *Current Neurology and Neuroscience Reports*, 15(9), 63. <https://doi.org/10.1007/s11910-015-0579-4>
- Neef, N. E., Büttfering, 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), Article e3002492. <https://doi.org/10.1371/journal.pbio.3002492>
- Neef, N. E., Korzeczek, A., Primafin, 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>
- Nora, A., Rinkinen, O., Renvall, H., Service, E., Arkkila, E., Smolander, S., Laasonen, M., & Salmelin, R.** (2024). Impaired cortical tracking of speech in children with developmental language disorder. *Journal of Neuroscience*, 44(22), Article e2048232024. <https://doi.org/10.1523/JNEUROSCI.2048-23.2024>
- Nuwer, M. R., Comi, G., Emerson, R., Fuglsang-Frederiksen, A., Guérit, J.-M., Hinrichs, H., Ikeda, A., Lucas, J. C. 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](https://doi.org/10.1016/S0013-4694(97)00106-5)
- Obleser, J., & Kayser, C.** (2019). Neural entrainment and attentional selection in the listening brain. *Trends in Cognitive Sciences*, 23(11), 913–926. <https://doi.org/10.1016/j.tics.2019.08.004>

- Oldfield, R. C.** (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia*, *9*(1), 97–113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4)
- 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*, 1–9. <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>
- Perron, M., Vuong, V., Grassi, M. W., Imran, A., & Alain, C.** (2024). Engagement of the speech motor system in challenging speech perception: Activation likelihood estimation meta-analyses. *Human Brain Mapping*, *45*(13), Article e70023. <https://doi.org/10.1002/hbm.70023>
- 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>
- Poeppl, D., & Assaneo, M. F.** (2020). Speech rhythms and their neural foundations. *Nature Reviews Neuroscience*, *21*(6), 322–334. <https://doi.org/10.1038/s41583-020-0304-4>
- Poeppl, D., & Teng, X.** (2020). Entrainment in human auditory cortex: Mechanism and functions. In B. Fritzsche (Ed.), *The senses: A comprehensive reference* (2nd ed., Vol. 2, pp. 63–76). Elsevier. <https://doi.org/10.1016/B978-0-12-805408-6.00018-X>
- Rauschecker, J. P.** (2012). Ventral and dorsal streams in the evolution of speech and language. *Frontiers in Evolutionary Neuroscience*, *4*, Article 18154. <https://doi.org/10.3389/FNEVO.2012.00007>
- Riecke, L., Formisano, E., Sorger, B., Başkent, D., & Gaudrain, E.** (2018). Neural entrainment to speech modulates speech intelligibility. *Current Biology*, *28*(2), 161–169.e5. <https://doi.org/10.1016/j.cub.2017.11.033>
- Riley, G. D.** (2009). *Stuttering Severity Instrument—Fourth Edition*. Pro-Ed.
- Rimmele, J. M., Morillon, B., Poeppl, 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>
- Scott, S. K., & Johnsrude, I. S.** (2003). The neuroanatomical and functional organization of speech perception. *Trends in Neurosciences*, *26*(2), 100–107. [https://doi.org/10.1016/S0166-2236\(02\)00037-1](https://doi.org/10.1016/S0166-2236(02)00037-1)
- 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>
- Smith, A., & Weber, C.** (2017). How stuttering develops: The multifactorial dynamic pathways theory. *Journal of Speech, Language, and Hearing Research*, *60*(9), 2483–2505. https://doi.org/10.1044/2017_JSLHR-S-16-0343
- 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](https://doi.org/10.1016/S0140-6736(02)09610-1)
- Stasak, B., Epps, J., Schatten, H. T., Miller, I. W., Provost, E. M., & Arme, M. F.** (2021). Read speech voice quality and disfluency in individuals with recent suicidal ideation or suicide attempt. *Speech Communication*, *132*, 10–20. <https://doi.org/10.1016/J.SPECOM.2021.05.004>
- 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*, Article 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>
- Westner, B. U., Dalal, S. S., Gramfort, A., Litvak, V., Mosher, J. C., Oostenveld, R., & Schoffelen, J. M.** (2022). A unified view on beamformers for M/EEG source reconstruction. *NeuroImage*, *246*, Article 118789. <https://doi.org/10.1016/J.NEUROIMAGE.2021.118789>
- Yairi, E., & Ambrose, N.** (2013). Epidemiology of stuttering: 21st century advances. *Journal of Fluency Disorders*, *38*(2), 66–87. <https://doi.org/10.1016/J.JFLUDIS.2012.11.002>
- 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>
- Zhang, N., Yin, Y., Jiang, Y., & Huang, C.** (2022). Reinvestigating the neural bases involved in speech production of stutterers: An ALE meta-analysis. *Brain Sciences*, *12*(8), Article 1030. <https://doi.org/10.3390/brainsci12081030>