

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

Simone Gastaldon <sup>1,2</sup>, Pierpaolo Busan <sup>3</sup>, Nicola Molinaro <sup>4,5</sup>, Mikel Lizarazu <sup>4</sup>

1) Department of Developmental and Social Psychology (DPSS), University of Padova, Padova, Italy

2) Padova Neuroscience Center (PNC), University of Padova, Padova, Italy

3) Department of Medical, Surgical and Health Sciences, University of Trieste, Trieste, Italy

4) Basque Center on Cognition, Brain and Language (BCBL), Donostia-San Sebastián, Spain

5) Ikerbasque, Basque Foundation for Science, Bilbao, Spain

## Corresponding author:

Simone Gastaldon

Dipartimento di Psicologia dello Sviluppo e della Socializzazione (DPSS)

Università di Padova

Via Venezia 8, 35131, Padova (PD), Italy

E-mail: simone.gastaldon@unipd.it

**Conflicts of interest:** None.

**CRedit statement:** Conceptualization: S.G., N.M., M.L., P.B. Methodology: M.L., N.M., S.G. Formal Analysis: M.L., S.G. Investigation: S.G. Data Curation: M.L. Writing – Original Draft: S.G., P.B., M.L. Writing – Review & Editing: S.G., P.B., N.M., M.L. Visualization: S.G., M.L. Supervision: N.M. Project Administration: S.G., M.L.

26 **Abstract**

27

28 **Purpose:** Investigate cortical tracking of speech (CTS) in adults who stutter (AWS) compared to  
29 typically fluent adults (TFA) to test the involvement of the speech-motor network in tracking  
30 rhythmic speech information.

31

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

36

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

44

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

50

51 **Keywords:** cortical tracking of speech; developmental stuttering; neural oscillations

## 52 **Introduction**

53

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

60 DS likely has a multifactorial origin, comprising motor, linguistic, emotional, neural and  
61 genetic factors (Smith & Weber, 2017). In particular, in recent years, genetic factors have been  
62 identified (Barnes et al., 2016; Benito-Aragón et al., 2020; Chow et al., 2020; Frigerio-Domingues &  
63 Drayna, 2017; Kang et al., 2010; Kang & Drayna, 2012; Kraft & Yairi, 2011), which may facilitate  
64 the appearance of atypical structure, function and connectivity patterns of the central nervous system  
65 (Alm, 2021a; Craig-McQuaide et al., 2014; Etchell et al., 2018), also expressed as altered  
66 sensorimotor brain rhythms (Etchell et al., 2016; Ghaderi et al., 2018; Jenson et al., 2018, 2020; Joos  
67 et al., 2014; Saltuklaroglu et al., 2017). Such atypical neural traits lead to impaired capacities in  
68 sensorimotor planning and execution of speech (Alm, 2021b; Chang et al., 2019). Deficits in internal  
69 timing and motor coordination seem to be at the core of DS (Alm, 2004), affecting wide neural  
70 systems and comprising regions such as the basal ganglia, the supplementary motor area, the inferior  
71 frontal cortex, and temporal regions (Busan, 2020; Busan et al., 2019; Craig-McQuaide et al., 2014;  
72 Etchell et al., 2018; Watkins et al., 2008). Within this network, disrupted auditory-motor interactions  
73 have been reported for DS, with difficulties in integrating auditory feedback during speaking  
74 potentially contributing to overt stuttering behaviors (Bradshaw et al., 2021; Chang et al., 2016; Daliri  
75 & Max, 2015, 2018; Halag-Milo et al., 2016; Hesse, 2023; Kim et al., 2020). Interestingly, disrupted  
76 sensorimotor function in DS has also been associated with impairments in aspects of speech  
77 perception, specifically weaker or less efficient predictive processing during spoken language  
78 comprehension (Gastaldon et al., 2023). To evaluate the role of inefficient sensorimotor function in  
79 DS during speech listening, this study considers cortical tracking of speech (CTS).

80 CTS, sometimes also referred to as “speech-brain entrainment” (Obleser & Kayser, 2019), is  
81 the temporal alignment of internal low frequency brain rhythms (delta – 0.1-3 Hz – and theta – 4-7  
82 Hz – frequency bands) to acoustic energy fluctuations (envelope amplitude) of the external speech  
83 signal. CTS is considered to be a valuable index reflecting the efficiency of neural processing of  
84 quasi-rhythmic components of speech, especially of prosodic (delta) and syllabic (theta) information  
85 (Assaneo & Poeppel, 2018; Molinaro & Lizarazu, 2018; Poeppel & Assaneo, 2020; Poeppel & Teng,

86 2020); disrupting such alignment can interfere with speech intelligibility (Riecke et al., 2018).  
87 Importantly, a growing body of evidence supports the view that frontal, motor and premotor regions  
88 modulate CTS in the auditory cortex in a top-down manner (Keitel et al., 2018; Park et al., 2015).  
89 Evidence also suggests that there is a preferred frequency range at which activity in the motor and  
90 auditory cortices are coupled during speech listening. Such range lies within the theta band,  
91 specifically between 3 and 5 Hz, with a peak at 4.5 Hz (Assaneo & Poeppel, 2018). This range has  
92 been associated with the rate of both production and perception of syllabic rhythm across languages  
93 (Ding et al., 2017, Poeppel & Assaneo, 2020). In this scenario, it has been proposed that the motor  
94 system generates temporal predictions via efferent motor signals, leading to phase-resetting in  
95 auditory cortices and hence to the optimization of auditory perception (Rimmele et al., 2018). Recent  
96 behavioral evidence supports this account: higher individual speech production rates (higher fluency)  
97 and stronger auditory-motor synchronization (alignment of speech production to an external rhythm)  
98 were associated with better performance in a speech comprehension task (Lubinus et al., 2023).

99 Importantly, altered brain processes related to CTS have been proposed as a risk factor for the  
100 appearance of developmental speech and/or language disorders (Lizarazu et al., 2015; Molinaro et  
101 al., 2016; Di Liberto et al., 2018; Lizarazu et al., 2021a; Nora et al., 2024). It has been proposed that  
102 such deficits are linked to abnormalities in the cortico-basal-thalamo-cortical circuitry involved in the  
103 processing of sensory cues (such as beats in music and/or linguistic meter in speech), thus playing a  
104 role in processing and predicting events in a sequence (Ladányi et al., 2020). This may also be the  
105 case for people who stutter, especially if auditory-motor coupling is a contributory factor to both DS  
106 and CTS. Crucially, no evidence is currently available for brain processes related to CTS in stuttering.  
107 Thus, a better understanding of these phenomena in DS should be useful for improving our  
108 comprehension of 1) neuro-pathological mechanisms related to stuttering (both in the contexts of  
109 speech production and perception), and 2) the neural mechanisms involved in typical speech  
110 perception and production (and in their possible mutual interactions).

111 To address both these issues, in the present study we investigated whether adults who stutter  
112 (AWS; stuttering onset during childhood and persisting into adulthood) show altered tracking of the  
113 speech signal when compared to typically fluent adults (TFA; no diagnosis of speech disorders).  
114 Specifically, we measured speech-brain coherence on electroencephalographic (EEG) data, both at  
115 the sensor and the neural source level, during sentence listening in conditions that either overtly  
116 recruited the articulatory system (completing the sentence by naming a picture; listening-for-  
117 speaking) or not (passively listening to the entire sentence; listening-only), in order to assess whether  
118 the upcoming involvement of the speech-motor network may have modulatory effects on CTS (see  
119 Figure 1 and Materials and Methods). While not directly simulating everyday dyadic conversations,

120 the listening-for-speaking task still implies an alertness of the speech-motor system, in addition to  
121 higher level processes such as anticipation and planning (Corps et al., 2018), similar to the demands  
122 of conversational and turn-taking settings. We expected to find group differences in a restricted range  
123 within the theta band, at which activity in the auditory and motor regions is supposed to be inherently  
124 coupled and to contribute to the tracking of syllabic rhythm.

125 Further insight can be obtained by studying how cortical regions interact with each other  
126 during speech listening. Standard neural models of speech processing describe the presence of a  
127 dorsal and a ventral cortical stream, jointly with a cortico-subcortical loop involving cerebellum,  
128 thalamus and basal ganglia (Friederici, 2012; Hickok & Poeppel, 2004, 2007; Kotz & Schwartz,  
129 2010; Scott & Johnsrude, 2003). While the ventral stream (comprising the auditory cortex, the  
130 anterior and posterior temporal lobe, and the inferior frontal cortex) is primarily tasked with mapping  
131 auditory information onto linguistic representations during speech comprehension, the dorsal stream  
132 (comprising the auditory cortex, inferior parietal regions, posterior temporal cortex, motor, premotor  
133 and supplementary motor regions, and the inferior frontal cortex) is concerned with mapping auditory  
134 and motor representations of linguistic sounds, primarily during speech production. However, the  
135 dorsal stream seems to be useful also during speech listening, especially in challenging listening  
136 situations (Skipper et al., 2017), thanks to the role of this network in providing temporal structure for  
137 processing, jointly with the basal ganglia (Kotz & Schwartz, 2010). This should be particularly true  
138 when specifically targeting the tracking of rhythmic information in the speech stream, which requires  
139 sensibility to temporal regularities (Ladányi et al., 2020), as is the case of the present study. For these  
140 reasons, we also analyzed brain connectivity by using partial directed coherence (PDC) to quantify  
141 directional neural interactions between brain regions implicated in sensorimotor and speech  
142 processing. This analysis allows us to investigate frequency-specific directional communication  
143 between cortical regions during envelope tracking. Previous studies that looked at frequency-specific  
144 brain connectivity during speech envelope tracking in various contexts highlighted a variety of  
145 connectivity patterns. Importantly, many of such patterns involved regions part of the dorsal stream  
146 in the delta and theta frequency ranges (e.g., Becker & Hervais-Adelman, 2023; Giordano et al., 2017;  
147 Hincapié Casas et al., 2021; Lizarazu et al., 2021a; Molinaro et al., 2016; Park et al., 2015), supporting  
148 the idea that such network is relevant for tracking rhythmic patterns in the speech stream.

149 Based on the models and previous findings on frequency-specific connectivity in envelope  
150 tracking, we hypothesize that neural connectivity is reduced in stuttering especially in the dorsal  
151 stream, suggested to be responsible for auditory-motor transformations. This hypothesis is also  
152 supported by evidence showing that white matter tracts considered to be part of the dorsal stream are  
153 usually altered in people who stutter (Kronfeld-Duenias et al., 2016; Neef et al., 2018, 2022; Sommer

154 et al., 2002; Watkins et al., 2008). In fact, DS is characterized by abnormal structural and functional  
155 connectivity of large neural circuits, especially involving sensorimotor and speech brain regions (for  
156 a recent review, see Etchell et al., 2018). For example, concerning measures of white matter integrity  
157 and efficiency, Sommer et al. (2002) showed that DS is characterized by lower fractional anisotropy  
158 in fibers below cortical regions representing larynx and tongue muscles in the left primary motor and  
159 somatosensory cortices, suggesting disconnection among speech/motor cortical areas. Watkins et al.  
160 (2008) found white matter deficits in regions underlying premotor cortices, suggesting impairments  
161 in connectivity with temporal regions, primary motor cortex, and inferior parietal cortex, thus  
162 weakening integration of speech and sensory feedback, as well as execution of articulatory  
163 movements in stuttering. Neef et al. (2018) disentangled the (adaptive or maladaptive) compensatory  
164 role of right hemisphere speech/motor regions in DS showing that stuttering severity correlates with  
165 connectivity of right frontal (hyperactive) regions, as well as with fractional anisotropy of the left  
166 superior longitudinal fascicle after speech therapy (Neef et al., 2022). When considering functional  
167 evidence (especially obtained from EEG data, as employed in this work), Busan et al. (2019) suggest  
168 that DS is characterized by a delay in recruiting left hemisphere fronto-parietal networks, followed  
169 by activation in homologous regions of the right hemisphere, thus suggesting a possible neural  
170 substrate for the emergence of dysfluencies. Along this line, Caruso et al. (2023) found a reduced  
171 connectivity of the left sensorimotor circuits in motor tasks in stuttering. More specifically, inter-  
172 hemispheric connectivity was weaker at lower frequencies (delta and theta range) and stronger in the  
173 beta band, suggesting that the right hemisphere might be recruited to support sensorimotor processing  
174 in DS and that an altered balance of the neural activity might be a fundamental aspect of stuttering.

175 All this considered, a better understanding of neural connectivity is fundamental also in the  
176 context of the present work, especially when hypothesizing the relevance of auditory-motor pathways  
177 in tracking rhythmic information during speech listening. However, we remain agnostic as to specific  
178 patterns (directionality) and potential additional group differences (e.g., stronger connectivity in  
179 AWS in other pathways, reflecting compensatory mechanisms). To this extent, the PDC analysis is  
180 partially theoretically driven (identify pathways compatible with a dorsal processing stream) and  
181 partially exploratory.

182 In summary, given the picture outlined above, we can hypothesize that:

183 1) CTS may be reduced in AWS relative to TFA, regardless of the listening condition (“listening-  
184 only” vs “listening-for-speaking”). Alternatively, differences may be detected only when listening is  
185 coupled with the upcoming necessity to overtly activate the speech-motor system, which is  
186 consequently kept in an “alert mode” in order to appropriately initiate speech (see “Sensor level  
187 analysis” section).

188 2) In AWS, reduced CTS may be found in speech-motor and premotor regions, in addition to auditory  
189 and associative regions (see “Source level analysis” section).

190 3) In AWS, regions that are considered to be part of the dorsal stream (inferior frontal cortex,  
191 premotor and supplementary motor regions, sensorimotor and temporo-parietal regions) may be  
192 communicating less efficiently with auditory regions during speech tracking, thus displaying reduced  
193 connectivity (see “Partial directed coherence (PDC) analysis”).

194

195

[FIGURE 1 HERE]

196

## 197 **Materials and Methods**

198

### 199 *Participants*

200

201 We analyzed CTS in a dataset collected for a previous study on spoken sentence processing in adults  
202 who stutter, which focused on different time-windows, used different analyses and had different aims  
203 (Gastaldon et al., 2023). The original study included 14 right-handed male adults who stutter (AWS)  
204 and 14 right-handed typically fluent male adults (TFA). The participants were matched for age and  
205 handedness. All participants were native speakers of Italian. The original study was approved by the  
206 Ethical Committee for Psychological Research of the University of Padova (protocol n. 3073) and  
207 conducted in accordance with the Declaration of Helsinki. We refer the reader to the original study  
208 for further details on AWS recruitment and assessment. Out of the 28 participants of the original  
209 study, four participants were excluded due to excessively noisy EEG data during sentence frame  
210 presentation for the analyses conducted here. The remaining 24 participants, 12 AWS and 12 TFA,  
211 were matched for age (AWS: mean age = 34.44, SD = 9.37; TFA: mean age = 33.42, SD = 8.94) and  
212 handedness (AWS: mean = 83.75, SD = 20.57; TFA: mean = 85.00, SD = 23.06), as assessed by  
213 means of the Edinburgh Handedness Inventory (Oldfield, 1971). From the original study we also  
214 retrieved data about the Stuttering Severity Index (SSI-4; Riley, 2009) of each AWS, which assigns  
215 an overall individual score by considering various characteristics of stuttering events and the  
216 occurrence of concomitant movements associated with disfluencies (e.g. oro-facial and limb  
217 movements) in both reading and spontaneous speech tasks. The higher the SSI-4 score, the more  
218 severe the stuttering. The final set was composed of 6 participants with very mild severity, 4 with  
219 mild and 2 with severe. Exploratory correlations between SSI-4 and CTS and connectivity data were  
220 performed (statistical threshold for explorative correlations:  $p \leq .01$ , two-tailed).

221

222 *Stimuli and procedure*

223

224 The stimuli were the same used in Gastaldon et al. (2020, 2023) (see the OSF repository for additional  
225 information on stimuli characteristics: <https://osf.io/tcbsh/>). They consisted of 256 sentence frames  
226 (sentences without the final word, ranging in duration from 1.55 to 3.54 s; mean duration = 2.39, SD  
227 = 0.4), which were paired with 128 target words and 128 b/w line pictures (124x124 pixels), such that  
228 each word and picture appeared twice, completing a high and low constraint sentence frame. Spoken  
229 stimuli were uttered by a female native Italian speaker, recorded and digitized at 44.1 kHz using  
230 Audacity®. Audio files (\*.wav) were also segmented using Audacity. During the task, participants  
231 listened to the sentence frames, then, after an 800 ms pause, either they heard a word (listening-only  
232 task) or had to produce it by naming a picture (listening-for-speaking task), in two distinct blocks,  
233 which were counterbalanced across participants. In the listening-only task, true/false comprehension  
234 questions were asked at the end of the trial in 20% of the trials, to maintain the participant's  
235 engagement (see also Figure 1). Due to the aims of the original study (Gastaldon et al., 2023), half  
236 the sentence frames induced high constraining contexts for the final word and half induced low  
237 constraining contexts, thus manipulating target word predictability. However, in order to allow for a  
238 better estimation of CTS and increase signal-to-noise ratio (SNR) and statistical power (see below),  
239 we did not divide the sentence frames into high vs low constraining contexts in the present analyses.  
240 We recognize that this may be a highly relevant variable that should be investigated in future studies  
241 (for a study in the normal population, see Molinaro et al., 2021); however, here we were limited in  
242 terms of SNR and number of trials. Note that groups did not differ in the number of EEG data  
243 segments coming from high and low constraining sentences in either task, thus excluding the  
244 possibility of any effect of this factor on any group differences (see Supplementary Table S2).  
245 Therefore, here we focused on the manipulation of task demands, which implied two different  
246 listening conditions: listening for comprehension (listening-only) or listening in order to complete the  
247 sentence as quickly as possible by naming a picture (listening-for-speaking). Participants sat in a  
248 dimly lit room in front of a computer screen. The experimental material was delivered through E-  
249 Prime 2.0 (Psychology Software Tools, Pittsburgh, PA). Auditory stimuli were presented through  
250 built-in speakers. Responses (picture naming and true/false answers) were collected via a microphone  
251 set in front of the participant. In the listening-for-speaking task, audio recording started at the onset  
252 of the picture to be named and lasted for 2 seconds. The experimental paradigm is exemplified in  
253 Figure 1. For further details on the experimental design, we refer the reader to the original study  
254 (Gastaldon et al., 2023).

255

256 ***EEG data acquisition and preprocessing***

257

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

265 The preprocessing pipeline for the present work was the following. Heartbeat and EOG  
266 artifacts were identified using independent component analysis (ICA) and subtracted from the  
267 recordings in a linear manner. The ICA decomposition was carried out using the Infomax algorithm  
268 implemented in the Fieldtrip toolbox (Oostenveld et al., 2011). Across participants, the number of  
269 heartbeat and ocular components that were removed varied from 1 to 4 and 1 to 3 components,  
270 respectively. Furthermore, trials were visually inspected to discard any remaining artifacts. Bad  
271 channels were substituted with interpolated values computed as the average of the neighboring  
272 electrodes obtained through the triangulation method implemented in Fieldtrip. A minimum of 75%  
273 artifact-free trials per participant was required for inclusion in subsequent analyses. As noted above,  
274 this led to the exclusion of two participants from each group, resulting in a final sample of 24  
275 participants (12 AWS and 12 TFA). In the case of TFA, an average of 4.89% (SD = 3.1) trials and  
276 5.23% (SD = 3.8) trials were excluded for comprehension and production tasks, respectively.  
277 Similarly, for AWS, an average of 5.4% (SD = 3.59) trials and 8.13% (SD = 3.86) trials were excluded  
278 for comprehension and production tasks, respectively. Importantly, no significant group or task  
279 differences were observed in the number of excluded trials (all  $T_s < 1.6$ , all  $p_s > .11$ , two-tailed t-  
280 test). EEG data and MATLAB scripts for the analyses described in the following sections are  
281 available on a dedicated OSF repository: <https://osf.io/7gpyb/>.

282

283 ***Cortical tracking of speech (CTS) analysis***

284

285 *Sensor level analysis*

286

287 Coherence measures the degree of phase synchronization between two signals in the frequency  
288 domain. For each participant and condition, we used coherence to quantify the cortical tracking of  
289 speech (CTS), which represents the coupling between the speech temporal envelope and cortical

290 oscillations. We obtained the envelope of the speech signal from the Hilbert transformed broadband  
291 stimulus waveform. According to previous research in speech processing we expected to find strong  
292 CTS in the low-frequency ( $< 10$  Hz) spectrum and in temporal sensors (Molinaro et al., 2016;  
293 Molinaro & Lizarazu, 2018; Lizarazu et al., 2021b; Ershaid et al., 2024; Issa et al., 2024). Therefore,  
294 we selected a set of 12 channels, evenly distributed to cover the temporal lobes of the brain –  
295 precisely, 6 channels allocated over the left hemisphere (C3, C5, CP3, CP5, FC3, FC5) and additional  
296 6 over the right hemisphere (C4, C6, CP4, CP6, FC4, FC6). Artifact-free trials were segmented into  
297 1-second windows with 50% overlap. Coherence was then calculated using the cross-spectral density  
298 of the FFT of the two signals (i.e., speech envelope and EEG data segments), normalized by the power  
299 spectrum of each signal. For each EEG sensor, coherence was calculated in the 1 – 15 Hz frequency  
300 band with 1 Hz (inverse of the segment duration) frequency resolution (Molinaro et al., 2016;  
301 Molinaro & Lizarazu, 2018). This procedure was followed for each participant and task/listening  
302 condition.

303 To estimate the coherence bias, the auditory envelopes were randomly shuffled across epochs  
304 for each participant, and coherence was recalculated in 100 permutations. The coherence data from  
305 the selected sensors of interest were separately averaged for each hemisphere and then transformed  
306 into z-scores using the mean and standard deviation derived from the 100 random EEG-audio pairings  
307 for those sensors. For each condition and frequency bin, z-score transformations were computed using  
308 the task-specific mean and standard deviation obtained from the random pairing dataset, and with an  
309 equal number of trials as the actual EEG-audio pairing dataset.

310 For the statistical analysis, we calculated the mean CTS values (z-scored coherence) within  
311 the theta band, specifically in the 3-5 Hz frequency range. We focused on this frequency range  
312 because of two specific reasons: 1) a peak is present in our auditory stimuli in the same frequency  
313 range, indicating syllabic rhythm (see Supplementary Figure 2), and 2) the existence of a frequency-  
314 restricted preference for the coupling between auditory and motor regions, as explained in the  
315 Introduction (see also Assaneo & Poeppel, 2018). To assess group differences in each task, we  
316 conducted an ANOVA on the z-transformed coherence values, with hemisphere (left vs right) as the  
317 within-subject factor and group (TFA vs AWS) as the between-subject factor (considering effects of  
318 main factors and their interaction; post-hoc analyses conducted using t-test; statistical threshold at  $p$   
319  $\leq .05$ , two-tailed).

320

321 *Source level analysis*

322

323 Coherence values were also estimated at the source level for each participant and condition in the  
324 theta band (3-5 Hz), where significant results were observed at the scalp level. For the source level  
325 analysis, we utilized a frequency-domain adaptive spatial filtering imaging of coherent sources  
326 algorithm (Gross et al., 2001), implemented in the Fieldtrip toolbox. To establish the spatial  
327 relationship between electrode positions of the participants (defined with a template electrode layout)  
328 and the cortical mesh, we employed a standard boundary element head model (BEM) extracted from  
329 the Montreal Neurological Institute (MNI) template. This BEM consists of three 3-D surfaces (skin,  
330 skull, brain) derived from the MNI dataset. The forward model was computed using an 8 mm grid  
331 encompassing the entire brain compartments of the BEM, representing various source positions. To  
332 perform source analysis, we constructed common space filters utilizing the leadfield of each grid  
333 point and the cross-spectral density matrix (CSD). The CSD matrices were computed within the theta  
334 (4 Hz with  $\pm 1$  Hz frequency smoothing) band by applying the fast Fourier transform to 1-second data  
335 segments in sliding windows shifting in 0.5 seconds steps. As anticipated, the selection of the theta  
336 range was based on the observation of group effects at the sensor level occurring specifically at this  
337 frequency. Beamformer coefficients were computed considering the dominant source direction within  
338 all voxels and a regularization factor of 7% was applied. The coherence for each source location was  
339 estimated using the EEG data and the spatial filter in the theta band. To ensure comparability of source  
340 coherence values across subjects, we normalized individual coherence brain maps. For this reason,  
341 the coherence at each source was converted to a z-score value by subtracting the mean coherence  
342 across all sources and dividing by the standard deviation across all sources. Successively, for each  
343 group and condition, z-scored source coherence values were projected on the brain surface mesh  
344 image `BrainMesh_ICBM152_smoothed` from `Surf Ice` (Version 12.1;  
345 <https://www.nitrc.org/projects/surfsice/>)

346 Finally, based on previous functional neural evidence on DS we selected five regions of  
347 interest (ROIs) from the Automatic Anatomical Labeling (AAL; Tzourio-Mazoyer et al., 2002). More  
348 specifically, ROIs were defined considering that stuttering mainly affects neural networks that are  
349 fundamental for sensorimotor processing, thus impairing speech planning, programming, and  
350 execution (compare with Chang et al., 2019). In this context, abnormal neural activity in areas such  
351 as the inferior frontal cortex, primary somato-motor regions, auditory cortex, supplementary motor  
352 area, premotor cortex, and associative regions (such as the parietal cortex) have been consistently  
353 reported as neural markers of DS (see Belyk et al., 2015, 2017; Brown et al., 2005; Budde et al., 2014;  
354 Busan, 2020; Busan et al., 2019; Chang & Guenther, 2020; Chang et al., 2019; Craig-McQuaide et  
355 al., 2014; Etchell et al., 2018; Ingham et al., 2012; Neef et al., 2015; Zhang et al., 2022). Therefore,  
356 within each cerebral hemisphere, we defined the subsequent “clusters” of brain regions of interest (as

357 shown in Figure 2): i) the inferior frontal gyrus (IFG, comprising the pars opercularis, pars  
 358 triangularis, and pars orbitalis), ii) the premotor and supplementary motor cortex (preM), iii) the  
 359 sensorimotor strip (SM, comprising the pre-central and post-central gyri), iv) the inferior parietal  
 360 lobule (IPL, comprising the supramarginal and angular gyri), and v) the superior temporal gyrus  
 361 (STG).

362 For each task, we employed the Wilcoxon ranked sum non-parametric test to assess group  
 363 differences on the mean of the z-scored coherence values within each ROI (statistical threshold at  $p$   
 364  $\leq .05$ , two-tailed).

365 [FIGURE 2 HERE]

366  
 367 *Partial directed coherence (PDC) analysis*

368  
 369 We employed partial directed coherence (PDC) to assess the causal connections between neural  
 370 activity associated with speech processing within our designated ROIs (IFG, preM, SM, IPL and  
 371 STG). After creating spatial filters, virtual time series in the source locations within the ROIs were  
 372 reconstructed by applying the respective spatial filter to the EEG sensor data filtered in the theta (3 -  
 373 5 Hz) band. Because ROIs typically comprise many point sources, we employed principal component  
 374 analysis (PCA) to identify the most representative time series within each ROI. To achieve this, we  
 375 conducted a PCA on all time-series within each ROI and selected the first principal vector, which  
 376 represented the distribution that explained most of the variance across all time-series that entered the  
 377 PCA. For each participant and task, we computed PDC between the representative time series in each  
 378 ROI. PDC is based on the Granger Causality principle (Granger, 1969; Seth et al., 2015) and on vector  
 379 autoregressive (VAR) modeling of the data. The VAR model of order  $p$  for a variable  $x$  is given by:

380  
 381 
$$x(t) = \sum_{r=1}^p a(r)s(t-r) + \varepsilon(t)$$

382  
 383 
$$(s_1(t) \cdots s_M(t)) = \sum_{r=1}^p a_r(s_1(k-r) \cdots s_M(k-r)) + (\varepsilon_1(t) \cdots \varepsilon_M(t))$$

384  
 385 where  $s(t) = (s_1(t), s_2(t), \dots, s_M(t))$  are the stationary M-dimensional simultaneously measured time  
 386 series in each ROI;  $a_r$  are the M x M coefficient matrices of the model; and  $(t)$  is a multivariate  
 387 Gaussian white noise process. In our case, M = 10 since we calculated the connectivity network

388 formed by five different ROIs. The model order  $p$  was selected with the Schwartz Information  
 389 Criterion. This criterion selects the model order that optimizes the goodness of fit of the model, while  
 390 introducing a penalty depending on the complexity of the model. In the frequency domain the version  
 391 of Granger-causality is given by:

392

393

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

394

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

397

398

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

399

400 The PDC provides a measure of the linear directional coupling strength of  $s_j$  on  $s_i$  at frequency  $f$   
 401 (theta). The PDC values vary between 0 (no directional coupling) and 1 (perfect directional coupling).  
 402 PDC analysis was performed using the Frequency-Domain Multivariate Analysis toolbox (FDMA,  
 403 Freiburg Center for Data Analysis and University of Freiburg, Germany), and the model order was  
 404 computed using algorithms developed in the Multivariate Autoregressive Model Fitting (ARfit)  
 405 software package (Schneider & Neumaier, 2001). To assess group differences, separately for each  
 406 task we used the Wilcoxon ranked sum non-parametric test on PDC values (statistical threshold at  $p$   
 407  $\leq .05$ , two-tailed).

408

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

410

411 For naming latencies (response times, RT), we took the data from Gastaldon et al. (2023), also  
 412 available here: <https://osf.io/5jkur/>. Here we summarize how latencies were derived in the original  
 413 study, but we refer the reader to the original article for additional details. To estimate naming times,  
 414 audio recordings (2 seconds \*.wav files starting at picture onset) were fed to Chronset (Roux et al.,  
 415 2017). Only correct responses were considered. Responses were coded as incorrect if they started  
 416 with hesitation sounds, if corrections were made during the response, or if the participant could not  
 417 produce enough of the target word in the 2-second recording (in order to be able to assess the  
 418 correctness of the response).

419 Statistical analyses were performed in R. Accuracy was analyzed with a generalized linear  
420 mixed-effects model (GLMM) with binomial distribution family. Group, lexical frequency of the  
421 target word (retrieved through PhonItalia; Goslin et al., 2014) and repetition (the same target picture  
422 was presented twice in the task, associated with two different sentence frames) were set as fixed  
423 effects, while participant and item as random intercepts. RTs were analyzed with a GLMM with  
424 gamma distribution family and identity link function. Group, lexical frequency of the target word and  
425 repetition were set as fixed effects, while participant and item as random intercepts. As explained  
426 above, we decided not to include sentence constraint as a factor here since for the coherence analysis  
427 (the main focus of the present work) we did not differentiate between the two conditions for  
428 methodological reasons. GLMM were fitted with the *lme4* package (Bates et al., 2015) and contrasts  
429 set to sum coding. Finally, as for SSI-4, RTs were correlated with CTS and connectivity data  
430 (statistical threshold at  $p \leq .01$ , two-tailed).

431

## 432 **Results**

433

### 434 *Naming (listening-for-speaking task)*

435

436 Accuracy and response times (RTs) are shown in Figure 3, while model summaries are reported in  
437 Table 1. Participants of both groups had a very high accuracy in producing the correct word (AWS:  
438 mean = 0.96, SD = 0.2; TFA: mean = 0.99, SD = 0.1; see Figure 3A). However, the model revealed  
439 a main effect of repetition (higher accuracy when the picture appeared for the second time) and a  
440 main effect of group, with AWS less accurate than TFA (see Table 1). Regarding response times,  
441 AWS were slower than TFA (AWS: mean = 771.19 ms, SD = 267.77; TFA: mean = 650.53 ms, SD  
442 = 219.98; see Figure 3B). The model revealed a main effect of repetition and, importantly, a main  
443 effect of group (see Table 1). To test the robustness of the results to possible outliers for accuracy,  
444 we re-run the analysis by excluding the AWS participant with accuracy = 0.84 (see Figure 3), and the  
445 results are still consistent (main effect of group:  $t = -2.35$ ,  $p = .019$ , 95% CI [0.33 - 0.91]).  
446 Speculatively, lower accuracy, rather than reflecting possible inefficiency in retrieving lexical items  
447 in AWS, is likely due to the limited time available for recording the responses (2 seconds after picture  
448 onset): sometimes, AWS may have provided the correct response outside this window, making it  
449 impossible to evaluate their response off-line, hence the reduced accuracy (i.e., fewer trials coded as  
450 correct). This interpretation is compatible with the fact that the accuracy outlier in the AWS group  
451 (accuracy = 0.84) is also the one with longest mean RTs for correct responses (RT = 1155 ms).

452

453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486

[FIGURE 3 HERE]

[TABLE 1 HERE]

### ***Sensor-level CTS***

Initially, we conducted an assessment of sensor-level cortical tracking of speech within the 1 - 15 Hz frequency range for each group (TFA and AWS) and task (listening-only and listening-for-speaking). Consistent with previous studies, we observed that during speech listening, CTS was highest in the theta (3-5 Hz) frequency band (Figure 4A) in bilateral fronto-central, central, and centro-parietal sensors (Figure 4B), consistent with the topography usually found in M/EEG studies on coherence as measure of CTS in the theta range (Destoky et al., 2019).

[FIGURE 4 HERE]

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  $p_s > .17$ ,  $\eta^2_s < 0.06$ ). However, we did observe a main effect of Group ( $F(1,22) = 4.07$ ,  $p = .05$ ,  $\eta^2 = 0.15$ ) in the CTS values for the listening-for-speaking task. Post-hoc tests showed that CTS was significantly higher in TFA compared to AWS ( $t = 2.02$ ,  $p = .05$ , 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 the listening-for-speaking tasks, frontal, temporal, and parietal cortical regions showed strong CTS (z-scored coherence values) in the theta band (Figure 5). Subsequently, we calculated the mean of the CTS values in each of the ROIs described in the Materials and Methods section: the inferior frontal gyrus (IFG), the premotor/supplementary motor cortex (preM), the sensorimotor strip (SM), the inferior parietal lobule (IPL), and the superior temporal gyrus (STG).

[FIGURE 5 HERE]

487 In line with the results observed at the sensor level, we found that the CTS values were significantly  
488 stronger for individuals with TFA compared to AWS in the left IPL ( $M_{TFA} = 1.92$ ,  $SD_{TFA} = 2.58$ ;  
489  $M_{AWS} = 0.22$ ,  $SD_{AWS} = 0.34$ ;  $p = .03$ ) and in the right preM regions ( $M_{TFA} = 0.07$ ,  $SD_{TFA} = 1.21$ ;  
490  $M_{AWS} = -0.74$ ,  $SD_{AWS} = 0.36$ ;  $p = .01$ ), only for the listening-for-speaking task (Figure 6). We did not  
491 find any group differences in the listening-only task (all  $p > .09$ ) (Supplementary Figure 1). No  
492 statistically significant correlations with RTs and SSI-4 were found.

493

494

[FIGURE 6 HERE]

495

### 496 *Source-level connectivity*

497

498 We used partial directed coherence (PDC) to assess causal functional connectivity during speech  
499 processing in the theta band (3 - 5 Hz) between different ROIs. For the listening-only task (Figure 7),  
500 we observed significantly higher connectivity in TFA compared to AWS from the left STG to the  
501 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  
502 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-  
503 for-speaking task (Figure 5), we observed significantly higher connectivity in TFA compared to AWS  
504 from the right STG to the left IPL ( $M_{TFA} = 0.05$ ,  $SD_{TFA} = 0.06$ ;  $M_{AWS} = 0.03$ ,  $SD_{AWS} = 0.02$ ;  $p = .05$ )  
505 and from the right STG to the left SM regions ( $M_{TFA} = 0.11$ ,  $SD_{TFA} = 0.06$ ;  $M_{AWS} = 0.06$ ,  $SD_{AWS} =$   
506  $0.04$ ;  $p = .03$ ).

507

508

[FIGURE 7 HERE]

509

510 Interestingly, when considering both groups together, a statistically significant negative  
511 correlation was found in the listening-for-speaking task between RTs and the connectivity from the  
512 right STG to the left SM cortex ( $r = -0.56$ ,  $p = .0048$ ): stronger directional connectivity between these  
513 regions is associated with faster response times (see Figure 8).

514

515

[FIGURE 8 HERE]

516

### 517 **Discussion**

518

519 In the present work, we analyzed cortical tracking of speech (CTS) in a group of adults who stutter  
520 (AWS) and a group of typically fluent adults (TFA). To investigate the role of the alertness state of

521 the speech-motor system in CTS, we analyzed two different listening situations: listening-only (no  
522 upcoming involvement of speech production) and listening-for-speaking (listen to an unfinished  
523 sentence and complete it by naming a picture; upcoming overt engagement of the speech-motor  
524 system). We observed reduced coherence in the theta range (3-5 Hz) in AWS relative to TFA in the  
525 listening-for-speaking task, both at the sensor (bilaterally around the temporal regions) and the neural  
526 source levels. More specifically, at source level, AWS showed lower CTS in the left inferior  
527 parietal/temporo-parietal cortex and in the right premotor and supplementary motor regions. Cortical  
528 connectivity measures in the theta range were differently modulated for AWS as well, with weaker  
529 connections in both listening conditions, indicating lower inter-hemispheric information exchange  
530 between frontal, auditory/temporal, and sensorimotor regions. Notably, in the listening-for-speaking  
531 task, we also found slower (speech) response times in AWS, and a significant negative correlation  
532 between RTs and connectivity from the right STG to the left SM cortex when considering all  
533 participants, reinforcing the arguments we lay out next.

534

### 535 *Cortical tracking of syllabic rhythm is reduced in Developmental Stuttering when listening for* 536 *speaking*

537

538 The listening-for-speaking condition in this study required speech listening to be interwoven with  
539 speech production, similar to turn-taking in conversational settings (Levinson, 2016). This entails the  
540 ability to efficiently time the transition between listening and speaking, and appropriately plan  
541 production initiation while still attending to speech. The present findings suggest that CTS in AWS  
542 is impaired especially in such situations. As highlighted in the Introduction, CTS is a neural index  
543 reflecting the alignment of the phase of (internal) brain frequencies to acoustic features of the speech  
544 signal (Assaneo & Poeppel, 2018; Poeppel & Assaneo, 2020; Poeppel & Teng, 2020). Crucially,  
545 coupling between oscillatory activity in auditory and speech-motor regions in a restricted frequency  
546 range within the theta band seems to support this process, specifically for the tracking of syllabic  
547 rhythm (Assaneo & Poeppel, 2018; Keitel et al., 2018; Morillon & Baillet, 2017; Park et al., 2015).  
548 In this study we found that in a population characterized by inefficient timing and implementation of  
549 speech-motor processes, i.e., adults who stutter (Alm, 2004, 2021b; Busan, 2020; Chang & Guenther,  
550 2020), CTS is also affected as a result of disruptions to auditory-motor coupling.

551 At the source level, in the listening-for-speaking task, we observed CTS reduction in the left  
552 inferior parietal cortex and in the right premotor and supplementary motor regions in AWS compared  
553 to TFA. All these regions are key cortical substrates for speech-motor coordination. The inferior  
554 parietal lobule (IPL), comprising the supramarginal gyrus and the angular gyrus, has been associated

555 with a variety of functions, including verbal working memory, auditory spatial localization,  
556 sensorimotor integration, semantic processing and action-motor control (Binder et al., 2009;  
557 Binkofski & Buccino, 2018; Bzdok et al., 2016; Shum et al., 2011). Importantly, this region and the  
558 partially overlapping (non-anatomically defined) temporo-parietal junction (TPJ; Igelström &  
559 Graziano, 2017) are nodes in many dual-route models of speech and auditory processing (Friederici,  
560 2012; Hickok et al., 2011; Hickok & Poeppel, 2004, 2007; Rauschecker, 2012). For instance, in  
561 Hickock and Poeppel’s model, the Sylvian Parietal Temporal (Spt) area (located between the inferior  
562 parietal lobule and the posterior part of the superior temporal gyrus, thus situated within the TPJ) is  
563 proposed to be an interface between auditory codes and motor programs supporting successful  
564 sensorimotor integration during speech production, instantiated in the dorsal pathway. The IPL is also  
565 key in neurocomputational models of speech production such as the DIVA/GODIVA models  
566 (Guenther, 2016), which propose that somatosensory error maps of the difference between intended  
567 and actual somatic states are computed in the IPL during speech production. Importantly, in the  
568 adjacent posterior STG/TPJ, auditory error maps are computed by comparing auditory feedback and  
569 predicted targets via motor efference copies (Guenther, 2016). Interestingly, even if not properly part  
570 of the “classical” cortico-basal-thalamo-cortical network involved in DS (Alm, 2004; Busan, 2020;  
571 Chang & Guenther, 2020; Craig-McQuaide et al., 2014), it is not uncommon for this region to be  
572 highlighted as part of a defective system in the brain of people who stutter (Busan et al., 2019; Neef  
573 et al., 2015; Yang et al., 2016).

574 On the other hand, the premotor ROI in our study encompasses the premotor cortex and the  
575 supplementary motor complex (supplementary motor area – SMA – and pre-SMA). In the speech-  
576 motor control literature, these regions have been highlighted in the composition and the timing of  
577 execution of speech-motor command sequences (Alario et al., 2006; Ghosh et al., 2008; Guenther,  
578 2016). More specifically, in the DIVA/GODIVA models, the SMA is responsible for the correct  
579 initiation of stored speech motor units, while the pre-SMA represents the global sequential structure  
580 of the syllables to be produced. On the other hand, these models propose that right hemisphere  
581 premotor regions may be a component of a feedback/control speech-motor network-(Bohland et al.,  
582 2010; Guenther, 2016; Tourville & Guenther, 2011; see Chang & Guenther, 2020; Civier et al., 2013  
583 for a perspective on DS). Notably, rhythm processing seems to be particularly reliant upon such  
584 cortical structures (together with subcortical regions), both in the speech and non-speech domains  
585 (Cannon & Patel, 2021; Fiveash et al., 2021; Kasdan et al., 2022; Ladányi et al, 2020). Additionally,  
586 the SMA has also been linked to the mediation of motor-sound representations in auditory prediction  
587 and speech imagery (Lima et al., 2016). Crucially, premotor and supplementary motor regions are  
588 among the regions that are found to be most dysfunctional in DS (Busan, 2020; Busan et al., 2019;

589 Chang & Guenther, 2020; Civier et al., 2013; Etchell et al., 2018). When related with present findings,  
590 this body of evidence is compatible with a key role of premotor/supplementary motor regions in  
591 tracking rhythmic information at the syllabic level during speech perception,—speculatively by  
592 transforming motor information into auditory templates for syllabic tracking.

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

601         The fact that we found differences in CTS within the theta range is also particularly interesting  
602 from a speech-motor impairment point of view. The theta rhythm has often been associated with  
603 syllabic grouping across languages, more specifically to the acoustic energy fluctuations of speech  
604 sound clusters organized around an energy peak (usually a vowel) (Strauß & Schwartz, 2017; see also  
605 Molinaro & Lizarazu, 2018; Poeppel & Assaneo, 2020). Indeed, a peak was present in the theta range  
606 in our audio stimuli (3-4 Hz), reflecting this acoustic property (see Supplementary Figure 2).  
607 Importantly, the syllable has been proposed to be an “interface” between the perceptual and the  
608 articulatory systems (Poeppel & Assaneo, 2020; Strauß & Schwartz, 2017). Articulators are  
609 biomechanically constrained as to the possible configurations they can produce and the speed at  
610 which they can be executed; the syllable represents the optimal motor-programming unit that the  
611 neural system can send to the motor system for execution (Guenther, 2016; Poeppel & Assaneo,  
612 2020). Crucially, individuals with DS seems to be impaired in the ability to automatically activate  
613 syllabic motor units associated with learned sound sequences via the basal ganglia motor loop  
614 connected to the pre-SMA and SMA (Alm, 2004, 2021a, 2021b; Busan, 2020; Chang & Guenther,  
615 2020; Civier et al., 2013).

616         Therefore, it appears that there is a circular relationship that, stemming from biomechanical  
617 articulation constraints via neural motor program units, leads to the acoustic - and hence perceptual -  
618 phenomenon of syllabic rhythm tracking (Poeppel & Assaneo, 2020; Strauß & Schwartz, 2017). We  
619 believe that this proposed circle of joint causes is closely related to the results obtained in the present  
620 study: the speech-motor production system is involved in tracking acoustic properties that arise from  
621 articulatory-motor constraints. When such a system is unstable (as in the case of DS), perceptual  
622 tracking is also less efficient, more noticeably when listening and speaking are interwoven and

623 partially overlapping neural resources are required, thus overburdening an already unstable system.  
624 Behaviorally, the presence of slower RTs in AWS further supports this view.

625

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

628

629 We found weaker connectivity patterns in AWS relative to TFA in both listening conditions. During  
630 the listening-only task, we found weaker directional connectivity from the left STG (auditory regions)  
631 to the right IFG and from the right IFG toward the left IPL/TPJ. In the listening-for-speaking task,  
632 we found weaker directional connectivity from the right STG to the left primary sensorimotor (SM)  
633 regions and to the left IPL/TPJ. While we do not interpret the different patterns across listening  
634 conditions, all the regions involved are consistent with a dorsal stream of processing (Friederici, 2012;  
635 Hickok & Poeppel, 2007), supporting the idea that auditory-motor mapping is important for cortical  
636 tracking of speech, at least of syllabic rhythm. This may be related to the nature of the syllable itself,  
637 representing the optimal motor unit for the human speech-motor system (Poeppel & Assaneo, 2020;  
638 Strauß & Schwartz, 2017). Reduced connectivity in AWS likely reflects lower availability of neural  
639 resources for information exchange between regions that are instrumental for auditory and motor  
640 processing and integration, compatible with recent proposals suggesting the presence of a general  
641 metabolic deficit in the stuttering brain (Alm, 2021a; see also Busan et al., 2019; Han et al., 2019;  
642 Maguire et al., 2021; Turk et al., 2021).

643 Importantly, in the listening-for-speaking task, we found a significant negative correlation  
644 between RTs and strength of right STG → left SM cortex connectivity: faster RTs were associated  
645 with increased connectivity between these regions. This may indicate that efficiently sending  
646 rhythmic auditory information to the primary sensorimotor cortex when speech listening has to be  
647 managed with (overt) upcoming speech-motor engagement facilitates speech production, possibly as  
648 a result of more efficient CTS and smoother transitioning between listening and speaking with  
649 concomitant speech planning. Note that this correlation, when explored separately for AWS and TFA,  
650 was not strongly evident in AWS ( $r = -0.44, p = .15$ ) but was present in TFA ( $r = -0.58, p = .045$ ; see  
651 Supplementary Table 1). As a further indication, we would also like to highlight that, albeit  
652 statistically not significant, an interesting trend was present in AWS when looking at this very same  
653 connectivity pattern and SSI-4, where a negative relation is found ( $r = -0.55, p = .06$ ; see  
654 Supplementary Table 1 and Supplementary Figure 3): higher SSI-4 scores - hence, more severe  
655 stuttering - were associated with weaker right STG → left SM cortex connectivity.

656 Together, these findings strongly suggest that stuttering may be associated with weaker  
657 connectivity between auditory and sensorimotor regions, which is fundamental for cortical tracking  
658 of syllabic units, in turn leading to poorer behavioral performance in terms of response times. This  
659 interpretation may be also compatible with proposals according to which an effective connection  
660 between these regions may be helpful for better managing (or “by-passing”) disfluencies, perhaps by  
661 exploiting rhythmic or tracked cues (Etchell et al., 2014). Consistent with this suggestive although  
662 marginal evidence, activity in right-hemispheric fronto-temporal regions is often reported as neural  
663 markers of stuttering trait and state (e.g., Belyk et al., 2015, 2017; Brown et al., 2005; Budde et al.,  
664 2014; Craig-McQuaide et al., 2014; Etchell et al., 2014, 2018; Ingham et al., 2012; Neef et al., 2015;  
665 Stasak et al., 2021), suggesting that they may have a role in compensatory (as well as in pathological)  
666 speech-motor programming and execution processes in AWS (Busan et al., 2019; Etchell et al., 2014;  
667 Neef et al., 2015, 2016, 2018b, 2023).

668

#### 669 *Significance of present outcomes for CTS and DS research*

670

671 The present findings may advance research on both CTS and DS. More specifically, they suggest that  
672 1) CTS requires neural resources that sustain sensorimotor processes for facilitating speech  
673 perception and intelligibility, 2) DS may lead to suboptimal CTS, especially when additional  
674 resources are needed for supporting concomitant speech preparation for upcoming production, and 3)  
675 DS not only impairs speech programming and production but is a more complex neurodevelopmental  
676 disorder. Further research should clarify the extent to which DS impacts CTS (and vice versa), how  
677 this might affect people’s everyday life, hence widening the scope of possible interventions for  
678 stuttering. This is especially important in light of recent evidence suggesting that auditory-motor  
679 coupling (and individual speech production rates) may explain performance in speech comprehension  
680 tasks (Lubinus et al., 2023). Less efficient CTS may be related to more effortful spoken language  
681 comprehension at a subtle level. This is in line with the results reported in Gastaldon et al. (2023):  
682 AWS seem less efficient at generating predictions during listening, hypothesized as a result of the  
683 inability to fully exploit their speech-motor network. Thus, further studies should investigate whether  
684 there is a causal link between CTS and specific processes of speech comprehension such as  
685 prediction, and how this causal chain may impact people with different speech and language deficits,  
686 especially in interactive contexts (see also Gastaldon et al., 2024 on the importance of studying  
687 atypical populations for a better understanding of predictive speech processing). In conclusion,  
688 research should move towards turn-based and conversational contexts (e.g., Jackson et al., 2021;

689 Weiss, 1995) to investigate possible subtle differences in how spoken language comprehension is  
690 achieved in the stuttering brain.

691

### 692 ***Limitations***

693

694 The study provides interesting new results, suggesting future venues for CTS and DS research;  
695 however, some limitations need to be taken into consideration.

696 First of all, sample sizes are small. The primary reason lies in the difficulty in recruiting AWS  
697 participants. This is a common problem when studying neurodevelopmental disorders at low  
698 incidence in the population, such as DS (Jones et al., 2002). To address this, in line with increasingly  
699 relevant Open Science practices, multi-lab projects can be an efficient way to overcome small *N*'s  
700 and to generalize or disconfirm results from individual underpowered studies, and to appropriately  
701 quantify effect sizes by means of meta-analyses (Heinrich & Knight, 2020; Lange, 2020; McShane  
702 et al., 2019). Note that, by making data available, we provide material for future meta-analyses and/or  
703 re-analyses, in the spirit of Open Science.

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

710 Another limitation regards localization of cortical regions. This limitation is common to all  
711 studies employing EEG. However, good estimates can still be obtained when using a sufficient  
712 number of electrodes covering all the scalp (such as 64 electrodes in the present work), by following  
713 standardized electrode placement, and by imposing reliable biophysical constraints to forward and  
714 inverse solutions (Lantz et al., 2003; Michel & Brunet, 2019; Michel et al., 2004; Westner et al.,  
715 2022). Future studies may employ higher density EEG systems or MEG, combined with individual  
716 structural scans, in order to provide a more accurate picture.

717

### 718 **Conclusions**

719

720 The present work suggests that CTS recruits (pre-)motor regions and regions responsible for  
721 sensorimotor integration, as well as auditory regions, supporting views proposing an interaction  
722 between these networks in speech/language perception (Pickering & Garrod, 2013; Skipper et al.,

723 2017), in addition to their instrumental role in orchestrating successful speech production (Guenther,  
724 2016; Hickok et al., 2011). CTS seems to work less efficiently in DS, especially when additional  
725 neural resources are needed for managing listening-for-speaking conditions, as usually happens in  
726 more ecological communicative situations (Neef & Chang, 2024). A better understanding of CTS  
727 processes in DS under various circumstances may be informative for improving rehabilitation  
728 solutions for stuttering.

729

## 730 **Acknowledgments**

731

732 Research was supported by the University of Padova, by the Basque Government through the BERC  
733 2018–2021 program, and by the Spanish State Research Agency through BCBL's Severo Ochoa  
734 excellence accreditation CEX2020-001010-S. SG was supported by a postdoctoral research grant  
735 funded by the Fondazione CARIPARO through the PHD@UNIPD call at the University of Padova  
736 (grant CUP\_C93C23003190005). NM was supported by the Spanish Ministry of Science, Innovation  
737 and University (grants RTI2018-096311-B-I00, PDC2022-133917-I00, PCI2022-135031-2,  
738 PID2022-136991NB-I00), the Agencia Estatal de Investigación (AEI), the Fondo Europeo de  
739 Desarrollo Regional (FEDER). ML was supported by the Ramón y Cajal programme of the Spanish  
740 Ministry of Science and Universities (MICIU) (grant RYC2022-035497-I). Funding sources had no  
741 role in the collection, analysis, and interpretation of data. We thank Caroline Handley for  
742 proofreading the manuscript.

743

## 744 **Data Availability Statement**

745

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

747

## 748 **References**

749

- 750 Alario, F. X., Chainay, H., Lehericy, S., & Cohen, L. (2006). The role of the supplementary motor area (SMA) in word  
751 production. *Brain Research, 1076*(1), 129–143. <https://doi.org/10.1016/j.brainres.2005.11.104>
- 752 Alm, P. A. (2004). Stuttering and the basal ganglia circuits: A critical review of possible relations. *Journal of*  
753 *Communication Disorders, 37*(4), 325–369. <https://doi.org/10.1016/j.jcomdis.2004.03.001>
- 754 Alm, P. A. (2021a). Stuttering: A Disorder of Energy Supply to Neurons? *Frontiers in Human Neuroscience, 0*, 289.  
755 <https://doi.org/10.3389/FNHUM.2021.662204>
- 756 Alm, P. A. (2021b). The Dopamine System and Automatization of Movement Sequences: A Review With Relevance  
757 for Speech and Stuttering. *Frontiers in Human Neuroscience, 15*, 663.  
758 <https://doi.org/10.3389/FNHUM.2021.661880/>
- 759 American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders: DSM-5™*, 5th ed. In

760 *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association.

761 <https://doi.org/10.1176/appi.books.9780890425596>

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

764 Barnes, T. D., Wozniak, D. F., Gutierrez, J., Han, T. U., Drayna, D., & Holy, T. E. (2016). A Mutation Associated with  
765 Stuttering Alters Mouse Pup Ultrasonic Vocalizations. *Current Biology*, 26(8), 1009–1018.

766 <https://doi.org/10.1016/J.CUB.2016.02.068>

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

769 Becker, R., & Hervais-Adelman, A. (2023). Individual theta-band cortical entrainment to speech in quiet predicts word-  
770 in-noise comprehension. *Cerebral Cortex Communications*, 4(1), 1–13.

771 <https://doi.org/10.1093/TEXCOM/TGAD001>

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

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

776 <https://doi.org/10.1111/EJN.13512>

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

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

782 <https://doi.org/10.1093/CERCOR/BHP055>

783 Binkofski, F., & Buccino, G. (2018). The role of the parietal cortex in sensorimotor transformations and action coding.  
784 In *Handbook of Clinical Neurology* (1st ed., Vol. 151, pp. 467–479). Elsevier B.V. [https://doi.org/10.1016/B978-](https://doi.org/10.1016/B978-0-444-63622-5.00024-3)  
785 [0-444-63622-5.00024-3](https://doi.org/10.1016/B978-0-444-63622-5.00024-3)

786 Bradshaw, A. R., Lametti, D. R., & McGettigan, C. (2021). The Role of Sensory Feedback in Developmental  
787 Stuttering: A Review. *Neurobiology of Language*, 2(2), 308–334. [https://doi.org/10.1162/nol\\_a\\_00036](https://doi.org/10.1162/nol_a_00036)

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

790 <https://doi.org/10.1002/hbm.20140>

791 Budde, K. S., Barron, D. S., & Fox, P. T. (2014). Stuttering, induced fluency, and natural fluency: A hierarchical series  
792 of activation likelihood estimation meta-analyses. *Brain and Language*, 139, 99–107.

793 <https://doi.org/10.1016/J.BANDL.2014.10.002>

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

796 Busan, P., Del Ben, G., Russo, L. R., Bernardini, S., Natarelli, G., Arcara, G., Manganotti, P., & Battaglini, P. P.  
797 (2019). Stuttering as a matter of delay in neural activation: A combined TMS/EEG study. *Clinical*  
798 *Neurophysiology*, 130(1), 61–76. <https://doi.org/10.1016/J.CLINPH.2018.10.005>

799 Bzdok, D., Hartwigsen, G., Reid, A., Laird, A. R., Fox, P. T., & Eickhoff, S. B. (2016). Left inferior parietal lobe  
800 engagement in social cognition and language. *Neuroscience and Biobehavioral Reviews*, 68, 319–334.

801 <https://doi.org/10.1016/j.neubiorev.2016.02.024>

802 Cannon, J. J., & Patel, A. D. (2021). How Beat Perception Co-opts Motor Neurophysiology. *Trends in Cognitive*

803 *Sciences*, 25(2), 137–150. <https://doi.org/10.1016/J.TICS.2020.11.002/>

804 Caruso, V. C., Wray, A. H., Lescht, E., & Chang, S.-E. (2023). Neural oscillatory activity and connectivity in children

805 who stutter during a non-speech motor task. *Journal of Neurodevelopmental Disorders*, 15(1), 40.

806 <https://doi.org/10.1186/s11689-023-09507-8>

807 Chang, S. E., Chow, H. M., Wieland, E. A., & McAuley, J. D. (2016). Relation between functional connectivity and

808 rhythm discrimination in children who do and do not stutter. *NeuroImage: Clinical*, 12, 442–450.

809 <https://doi.org/10.1016/J.NICL.2016.08.021>

810 Chang, S. E., Garnett, E. O., Etchell, A., & Chow, H. M. (2019). Functional and Neuroanatomical Bases of

811 Developmental Stuttering: Current Insights. *The Neuroscientist*, 25(6), 566–582.

812 <https://doi.org/10.1177/1073858418803594>

813 Chang, S. E., & Guenther, F. H. (2020). Involvement of the Cortico-Basal Ganglia-Thalamocortical Loop in

814 Developmental Stuttering. *Frontiers in Psychology*, 10, 3088. <https://doi.org/10.3389/FPSYG.2019.03088/>

815 Chow, H. M., Garnett, E. O., Li, H., Etchell, A., Sepulcre, J., Drayna, D., Chugani, D. C., & Chang, S.-E. (2020).

816 Linking lysosomal enzyme targeting genes and energy metabolism with altered gray matter volume in children

817 with persistent stuttering. *Neurobiology of Language*, 1–40. [https://doi.org/10.1162/nol\\_a\\_00017](https://doi.org/10.1162/nol_a_00017)

818 Civier, O., Bullock, D., Max, L., & Guenther, F. H. (2013). Computational modeling of stuttering caused by

819 impairments in a basal ganglia thalamo-cortical circuit involved in syllable selection and initiation. *Brain and*

820 *Language*, 126(3), 263–278. <https://doi.org/10.1016/j.bandl.2013.05.016>

821 Corps, R. E., Gambi, C., & Pickering, M. J. (2018). Coordinating Utterances During Turn-Taking: The Role of

822 Prediction, Response Preparation, and Articulation. *Discourse Processes*, 55(2), 230–240.

823 <https://doi.org/10.1080/0163853X.2017.1330031>

824 Craig-McQuaide, A., Akram, H., Zrinzo, L., & Tripoliti, E. (2014). A review of brain circuitries involved in stuttering.

825 *Frontiers in Human Neuroscience*, 8(November), 1–20. <https://doi.org/10.3389/fnhum.2014.00884>

826 Craig, A., Blumgart, E., & Tran, Y. (2009). The impact of stuttering on the quality of life in adults who stutter. *Journal*

827 *of Fluency Disorders*, 34(2), 61–71. <https://doi.org/10.1016/J.JFLUDIS.2009.05.002>

828 Daliri, A., & Max, L. (2015). Modulation of auditory processing during speech movement planning is limited in adults

829 who stutter. *Brain and Language*, 143, 59–68. <https://doi.org/10.1016/J.BANDL.2015.03.002>

830 Daliri, A., & Max, L. (2018). Stuttering adults' lack of pre-speech auditory modulation normalizes when speaking with

831 delayed auditory feedback. *Cortex*, 99, 55–68. <https://doi.org/10.1016/J.CORTEX.2017.10.019>

832 Destoky, F., Philippe, M., Bertels, J., Verhasselt, M., Coquelet, N., Vander Ghinst, M., Wens, V., De Tiège, X., &

833 Bourguignon, M. (2019). Comparing the potential of MEG and EEG to uncover brain tracking of speech temporal

834 envelope. *NeuroImage*, 184, 201–213. <https://doi.org/10.1016/J.NEUROIMAGE.2018.09.006>

835 Di Liberto, G. M., Peter, V., Kalashnikova, M., Goswami, U., Burnham, D., & Lalor, E. C. (2018). Atypical cortical

836 entrainment to speech in the right hemisphere underpins phonemic deficits in dyslexia. *NeuroImage*, 175, 70–79.

837 <https://doi.org/10.1016/J.NEUROIMAGE.2018.03.072>

838 Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., & Poeppel, D. (2017). Temporal modulations in speech and music.

839 *Neuroscience & Biobehavioral Reviews*, 81, 181–187. <https://doi.org/10.1016/J.NEUBIOREV.2017.02.011>

840 Ershaid, H., Lizarazu, M., McLaughlin, D., Cooke, M., Simantiraki, O., Koutsogiannaki, M., & Lallier, M. (2024).

841 Contributions of listening effort and intelligibility to cortical tracking of speech in adverse listening conditions.

842 *Cortex*, 172, 54–71. <https://doi.org/10.1016/j.cortex.2023.11.018>

843 Etchell, A., Adhikari, A., Weinberg, L. S., Choo, A. L., Garnett, E. O., Chow, H. M., & Chang, S.-E. (2018). A  
844 systematic literature review of sex differences in childhood language and brain development. *Neuropsychologia*,  
845 114, 19–31. <https://doi.org/10.1016/j.neuropsychologia.2018.04.011>

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

849 Etchell, A. C., Johnson, B. W., & Sowman, P. F. (2014). Behavioral and multimodal neuroimaging evidence for a  
850 deficit in brain timing networks in stuttering: A hypothesis and theory. *Frontiers in Human Neuroscience*,  
851 8(JUNE), 467. <https://doi.org/10.3389/fnhum.2014.00467>

852 Etchell, A. C., Ryan, M., Martin, E., Johnson, B. W., & Sowman, P. F. (2016). Abnormal time course of low beta  
853 modulation in non-fluent preschool children: A magnetoencephalographic study of rhythm tracking. *NeuroImage*,  
854 125, 953–963. <https://doi.org/10.1016/j.neuroimage.2015.10.086>

855 Fiveash, A., Bedoin, N., Gordon, R. L., & Tillmann, B. (2021). Processing Rhythm In Speech And Music: Shared  
856 Mechanisms And Implications For Developmental Speech And Language Disorders. *Neuropsychology*, 35(8),  
857 771–791. <https://doi.org/10.1037/neu0000766>

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

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

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

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

868 Gastaldon, S., Bonfiglio, N., Vespignani, F., & Peressotti, F. (2024). Predictive language processing: integrating  
869 comprehension and production, and what atypical populations can tell us. *Frontiers in Psychology*, 15, 1369177.  
870 <https://doi.org/10.3389/fpsyg.2024.1369177>

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

873 Ghosh, S. S., Tourville, J. A., & Guenther, F. H. (2008). A neuroimaging study of premotor lateralization and cerebellar  
874 involvement in the production of phonemes and syllables. *Journal of Speech, Language, and Hearing Research*,  
875 51(5), 1183–1202. [https://doi.org/10.1044/1092-4388\(2008/07-0119\)](https://doi.org/10.1044/1092-4388(2008/07-0119))

876 Giordano, B. L., Ince, R. A. A., Gross, J., Schyns, P. G., Panzeri, S., & Kayser, C. (2017). Contributions of local speech  
877 encoding and functional connectivity to audio-visual speech perception. *ELife*, 6.  
878 <https://doi.org/10.7554/ELIFE.24763>

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

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

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

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

888 Halag-Milo, T., Stoppelman, N., Kronfeld-Duenias, V., Civier, O., Amir, O., Ezrati-Vinacour, R., & Ben-Shachar, M.  
889 (2016). Beyond production: Brain responses during speech perception in adults who stutter. *NeuroImage:*  
890 *Clinical*, 11, 328–338. <https://doi.org/10.1016/j.nicl.2016.02.017>

891 Han, T. U., Root, J., Reyes, L. D., Huchinson, E. B., Du Hoffmann, J., Lee, W. S., Barnes, T. D., & Drayna, D. (2019).  
892 Human GNPTAB stuttering mutations engineered into mice cause vocalization deficits and astrocyte pathology in  
893 the corpus callosum. *Proceedings of the National Academy of Sciences of the United States of America*, 116(35),  
894 17515–17524. <https://doi.org/10.1073/PNAS.1901480116/>

895 Heinrich, A., & Knight, S. (2020). Reproducibility in Cognitive Hearing Research: Theoretical Considerations and  
896 Their Practical Application in Multi-Lab Studies. *Frontiers in Psychology*, 11, 500520.  
897 <https://doi.org/10.3389/fpsyg.2020.01590>

898 Hesse, T. (2023). Developmental stuttering may be caused by insufficient processing of auditory feedback. *Medical*  
899 *Hypotheses*, 180, 111166. <https://doi.org/10.1016/J.MEHY.2023.111166>

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

902 Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional  
903 anatomy of language. *Cognition*, 92(1–2), 67–99. <https://doi.org/10.1016/j.cognition.2003.10.011>

904 Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8(5),  
905 393–402. <https://doi.org/10.1038/nrn2113>

906 Hincapié Casas, A. S., Lajnef, T., Pascarella, A., Guiraud-Vinatea, H., Laaksonen, H., Bayle, D., Jerbi, K., &  
907 Boulenger, V. (2021). Neural oscillations track natural but not artificial fast speech: Novel insights from speech-  
908 brain coupling using MEG. *NeuroImage*, 244, 118577. <https://doi.org/10.1016/J.NEUROIMAGE.2021.118577>

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

911 Ingham, R. J., Grafton, S. T., Bothe, A. K., & Ingham, J. C. (2012). Brain activity in adults who stutter: Similarities  
912 across speaking tasks and correlations with stuttering frequency and speaking rate. *Brain and Language*, 122(1),  
913 11–24. <https://doi.org/10.1016/j.bandl.2012.04.002>

914 Issa, M. F., Khan, I., Ruzzoli, M., Molinaro, N., & Lizarazu, M. (2024). On the speech envelope in the cortical tracking  
915 of speech. *NeuroImage*, 297, 120675. <https://doi.org/10.1016/j.neuroimage.2024.120675>

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

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

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

- 924 Jones, M., Gebiski, V., Onslow, M., & Packman, A. (2002). Statistical Power in Stuttering Research. *Journal of Speech,*  
 925 *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))
- 926 Joos, K., Ridder, D. De, Boey, R. A., & Vanneste, S. (2014). Functional connectivity changes in adults with  
 927 developmental stuttering: A preliminary study using quantitative electro-encephalography. *Frontiers in Human*  
 928 *Neuroscience*, 8(OCT). <https://doi.org/10.3389/fnhum.2014.00783>
- 929 Kang, C., & Drayna, D. (2012). A role for inherited metabolic deficits in persistent developmental stuttering. *Molecular*  
 930 *Genetics and Metabolism*, 107(3), 276–280. <https://doi.org/10.1016/J.YMGME.2012.07.020>
- 931 Kang, C., Riazuddin, S., Mundorff, J., Krasnewich, D., Friedman, P., Mullikin, J. C., & Drayna, D. (2010). Mutations in  
 932 the Lysosomal Enzyme–Targeting Pathway and Persistent Stuttering. *New England Journal of Medicine*, 362(8),  
 933 677–685. <https://doi.org/10.1056/NEJMOA0902630>
- 934 Kasdan, A. V., Burgess, A. N., Pizzagalli, F., Scartozzi, A., Chern, A., Kotz, S. A., Wilson, S. M., & Gordon, R. L.  
 935 (2022). Identifying a brain network for musical rhythm: A functional neuroimaging meta-analysis and systematic  
 936 review. *Neuroscience & Biobehavioral Reviews*, 136, 104588.  
 937 <https://doi.org/10.1016/J.NEUBIOREV.2022.104588>
- 938 Keitel, A., Gross, J., & Kayser, C. (2018). Perceptually relevant speech tracking in auditory and motor cortex reflects  
 939 distinct linguistic features. *PLOS Biology*, 16(3), e2004473. <https://doi.org/10.1371/journal.pbio.2004473>
- 940 Kim, K. S., Daliri, A., Flanagan, J. R., & Max, L. (2020). Dissociated Development of Speech and Limb Sensorimotor  
 941 Learning in Stuttering: Speech Auditory-motor Learning is Impaired in Both Children and Adults Who Stutter.  
 942 *Neuroscience*, 451, 1–21. <https://doi.org/10.1016/J.NEUROSCIENCE.2020.10.014>
- 943 Kotz, S. A., & Schwartze, M. (2010). Cortical speech processing unplugged: A timely subcortico-cortical framework.  
 944 *Trends in Cognitive Sciences*, 14(9), 392–399. <https://doi.org/10.1016/j.tics.2010.06.005>
- 945 Kraft, S. J., & Yairi, E. (2011). Genetic Bases of Stuttering: The State of the Art, 2011. *Folia Phoniatrica et*  
 946 *Logopaedica*, 64(1), 34–47. <https://doi.org/10.1159/000331073>
- 947 Kronfeld-Duenias, V., Amir, O., Ezrati-Vinacour, R., Civier, O., & Ben-Shachar, M. (2016). Dorsal and ventral  
 948 language pathways in persistent developmental stuttering. *Cortex*, 81, 79–92.  
 949 <https://doi.org/10.1016/j.cortex.2016.04.001>
- 950 Ladányi, E., Persici, V., Fiveash, A., Tillmann, B., & Gordon, R. L. (2020). Is atypical rhythm a risk factor for  
 951 developmental speech and language disorders? *Wiley Interdisciplinary Reviews: Cognitive Science, March*,  
 952 e1528. <https://doi.org/10.1002/wcs.152>
- 953 Lange, F. (2020). Are Difficult-To-Study Populations too Difficult to Study in a Reliable Way? *European Psychologist*,  
 954 25(1), 41–50. <https://doi.org/10.1027/1016-9040/a000384>
- 955 Lantz, G., Grave de Peralta, R., Spinelli, L., Seeck, M., & Michel, C. M. (2003). Epileptic source localization with high  
 956 density EEG: how many electrodes are needed? *Clinical Neurophysiology*, 114(1), 63–69.  
 957 [https://doi.org/10.1016/S1388-2457\(02\)00337-1](https://doi.org/10.1016/S1388-2457(02)00337-1)
- 958 Levinson, S. C. (2016). Turn-taking in Human Communication - Origins and Implications for Language Processing.  
 959 *Trends in Cognitive Sciences*, 20(1), 6–14. <https://doi.org/10.1016/j.tics.2015.10.010>
- 960 Lima, C. F., Krishnan, S., & Scott, S. K. (2016). Roles of Supplementary Motor Areas in Auditory Processing and  
 961 Auditory Imagery. *Trends in Neurosciences*, 39(8), 527–542. <https://doi.org/10.1016/j.tins.2016.06.003>
- 962 Lizarazu, M., Lallier, M., Molinaro, N., Bourguignon, M., Paz-Alonso, P. M., Lerma-Usabiaga, G., & Carreiras, M.  
 963 (2015). Developmental evaluation of atypical auditory sampling in dyslexia: Functional and structural evidence.  
 964 *Human Brain Mapping*, 36(12), 4986–5002. <https://doi.org/10.1002/HBM.22986>

- 965 Lizarazu, M., Lallier, M., Bourguignon, M., Carreiras, M., & Molinaro, N. (2021a). Impaired neural response to speech  
 966 edges in dyslexia. *Cortex*, *135*, 207-218. <https://doi.org/10.1016/j.cortex.2020.09.033>
- 967 Lizarazu, M., Scotto di Covella, L., van Wassenhove, V., Rivière, D., Mizzi, R., Lehongre, K., Hertz-Pannier, L., &  
 968 Ramus, F. (2021b). Neural entrainment to speech and nonspeech in dyslexia: Conceptual replication and  
 969 extension of previous investigations. *Cortex*, *137*, 160–178. <https://doi.org/10.1016/j.cortex.2020.12.024>
- 970 Lubinus, C., Keitel, A., Obleser, J., Poeppel, D., & Rimmele, J. M. (2023). Explaining flexible continuous speech  
 971 comprehension from individual motor rhythms. *Proceedings of the Royal Society B*, *290*(1994).  
 972 <https://doi.org/10.1098/RSPB.2022.2410>
- 973 Maguire, G. A., Yoo, B. R., & SheikhBahaei, S. (2021). Investigation of Risperidone Treatment Associated With  
 974 Enhanced Brain Activity in Patients Who Stutter. *Frontiers in Neuroscience*, *15*, 598949.  
 975 <https://doi.org/10.3389/FNINS.2021.598949/>
- 976 McShane, B. B., Tackett, J. L., Böckenholt, U., & Gelman, A. (2019). Large-Scale Replication Projects in  
 977 Contemporary Psychological Research. *The American Statistician*, *73*(sup1), 99–105.  
 978 <https://doi.org/10.1080/00031305.2018.1505655>
- 979 Michel, C.M., Murray, M.M., Lantz, G., Gonzalez, S., Spinelli, L., & Grave de Peralta, R. (2004). EEG source imaging.  
 980 *Clinical Neurophysiology* *115*(10), 2195-2222. <https://doi.org/10.1016/j.clinph.2004.06.001>
- 981 Michel, C. M., & Brunet, D. (2019). EEG Source Imaging: A Practical Review of the Analysis Steps. *Frontiers in*  
 982 *Neurology*, *10*, 325. <https://doi.org/10.3389/fneur.2019.00325>
- 983 Molinaro, N., Lizarazu, M., Lallier, M., Bourguignon, M., & Carreiras, M. (2016). Out-of-synchrony speech  
 984 entrainment in developmental dyslexia. *Human Brain Mapping*, *37*(8), 2767–2783.  
 985 <https://doi.org/10.1002/hbm.23206>
- 986 Molinaro, N., & Lizarazu, M. (2018). Delta (but not theta)-band cortical entrainment involves speech-specific  
 987 processing. *European Journal of Neuroscience*, *48*(7), 2642–2650. <https://doi.org/10.1111/ejn.13811>
- 988 Molinaro, N., Lizarazu, M., Baldin, V., Pérez-Navarro, J., Lallier, M., & Ríos-López, P. (2021). Speech-brain phase  
 989 coupling is enhanced in low contextual semantic predictability conditions. *Neuropsychologia*, 107830.  
 990 <https://doi.org/10.1016/j.neuropsychologia.2021.107830>
- 991 Morillon, B., & Baillet, S. (2017). Motor origin of temporal predictions in auditory attention. *Proceedings of the*  
 992 *National Academy of Sciences*, *114*(42), E8913–E8921. <https://doi.org/10.1073/pnas.1705373114>
- 993 Nang, C., Hersh, D., Milton, K., & Lau, S. R. (2018). The Impact of Stuttering on Development of Self-Identity,  
 994 Relationships, and Quality of Life in Women Who Stutter. *American Journal of Speech-Language Pathology*,  
 995 *27*(3S), 1244–1258. [https://doi.org/10.1044/2018\\_AJSLP-ODC11-17-0201](https://doi.org/10.1044/2018_AJSLP-ODC11-17-0201)
- 996 Neef, N. E., Angstadt, M., Koenraads, S. P. C., & Chang, S. E. (2023). Dissecting structural connectivity of the left and  
 997 right inferior frontal cortex in children who stutter. *Cerebral Cortex*, *33*(7), 4085–4100.  
 998 <https://doi.org/10.1093/CERCOR/BHAC328>
- 999 Neef, N. E., Anwander, A., Bütfering, C., Schmidt-Samoa, C., Friederici, A. D., Paulus, W., & Sommer, M. (2018).  
 1000 Structural connectivity of right frontal hyperactive areas scales with stuttering severity. *Brain*, *141*(1), 191–204.  
 1001 <https://doi.org/10.1093/brain/awx316>
- 1002 Neef, N. E., Anwander, A., & Friederici, A. D. (2015). The Neurobiological Grounding of Persistent Stuttering: from  
 1003 Structure to Function. *Current Neurology and Neuroscience Reports*, *15*(9), 1–11. [https://doi.org/10.1007/s11910-](https://doi.org/10.1007/s11910-015-0579-4)  
 1004 [015-0579-4](https://doi.org/10.1007/s11910-015-0579-4)
- 1005 Neef, N. E., Bütfering, C., Anwander, A., Friederici, A. D., Paulus, W., & Sommer, M. (2016). Left posterior-dorsal

1006 area 44 couples with parietal areas to promote speech fluency, while right area 44 activity promotes the stopping  
1007 of motor responses. *NeuroImage*, 142, 628–644. <https://doi.org/10.1016/J.NEUROIMAGE.2016.08.030>

1008 Neef, N. E., & Chang, S.-E. (2024). Knowns and unknowns about the neurobiology of stuttering. *PLOS Biology*, 22(2),  
1009 e3002492. <https://doi.org/10.1371/journal.pbio.3002492>

1010 Neef, N. E., Korzeczek, A., Primašin, A., Wolff von Gudenberg, A., Dechent, P., Riedel, C. H., Paulus, W., & Sommer,  
1011 M. (2022). White matter tract strength correlates with therapy outcome in persistent developmental stuttering.  
1012 *Human Brain Mapping*, 43(11), 3357–3374. <https://doi.org/10.1002/HBM.25853>

1013 Nora, A., Rinkinen, O., Renvall, H., Service, E., Arkkila, E., Smolander, S., Laasonen, M., & Salmelin, R. (2024).  
1014 Impaired Cortical Tracking of Speech in Children with Developmental Language Disorder. *Journal of*  
1015 *Neuroscience*, 44(22). <https://doi.org/10.1523/JNEUROSCI.2048-23.2024>

1016 Nuwer, M. R., Comi, G., Emerson, R., Fuglsang-Frederiksen, A., Guérit, J.-M., Hinrichs, H., Ikeda, A., Jose C. Luccas,  
1017 F., & Rappelsburger, P. (1998). IFCN standards for digital recording of clinical EEG. *Electroencephalography*  
1018 *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)

1019 Obleser, J., & Kayser, C. (2019). Neural Entrainment and Attentional Selection in the Listening Brain. *Trends in*  
1020 *Cognitive Sciences*, 23(11), 913–926. <https://doi.org/10.1016/j.tics.2019.08.004>

1021 Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1),  
1022 97–113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4)

1023 Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis  
1024 of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, 2011.  
1025 <https://doi.org/10.1155/2011/156869>

1026 Park, H., Ince, R. A. A., Schyns, P. G., Thut, G., & Gross, J. (2015). Frontal Top-Down Signals Increase Coupling of  
1027 Auditory Low-Frequency Oscillations to Continuous Speech in Human Listeners. *Current Biology*, 25(12), 1649–  
1028 1653. <https://doi.org/10.1016/J.CUB.2015.04.049>

1029 Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and*  
1030 *Brain Sciences*, 36(4), 329–347. <https://doi.org/10.1017/S0140525X12001495>

1031 Poeppel, D., & Assaneo, M. F. (2020). Speech rhythms and their neural foundations. *Nature Reviews Neuroscience*, 1–  
1032 13. <https://doi.org/10.1038/s41583-020-0304-4>

1033 Poeppel, D., & Teng, X. (2020). Entrainment in Human Auditory Cortex: Mechanism and Functions. In B. Fritzsche  
1034 (Ed.), *The Senses: A Comprehensive Reference (2nd ed.)*, Vol. 2, pp. 63–76. Elsevier.  
1035 <https://doi.org/10.1016/B978-0-12-805408-6.00018-X>

1036 Rauschecker, J. P. (2012). Ventral and dorsal streams in the evolution of speech and language. *Frontiers in*  
1037 *Evolutionary Neuroscience*, 4(MAY), 18154. <https://doi.org/10.3389/FNEVO.2012.00007>

1038 Riecke, L., Formisano, E., Sorger, B., Başkent, D., & Gaudrain, E. (2018). Neural Entrainment to Speech Modulates  
1039 Speech Intelligibility. *Current Biology*, 28(2), 161-169.e5. <https://doi.org/10.1016/j.cub.2017.11.033>

1040 Riley, G. D. (2009). *The stuttering severity instrument for adults and children (SSI-4)*. 4th Edition. PRO-ED Inc.

1041 Rimmele, J. M., Morillon, B., Poeppel, D., & Arnal, L. H. (2018). Proactive Sensing of Periodic and Aperiodic  
1042 Auditory Patterns. *Trends in Cognitive Sciences*, 22(10), 870–882. <https://doi.org/10.1016/J.TICS.2018.08.003>

1043 Roux, F., Armstrong, B. C., & Carreiras, M. (2017). Chronset: An automated tool for detecting speech onset. *Behavior*  
1044 *Research Methods*, 49(5), 1864–1881. <https://doi.org/10.3758/s13428-016-0830-1>

1045 Saltuklaroglu, T., Harkrider, A. W., Thornton, D., Jenson, D., & Kittilstved, T. (2017). EEG Mu ( $\mu$ ) rhythm spectra and  
1046 oscillatory activity differentiate stuttering from non-stuttering adults. *NeuroImage*, 153, 232–245.

- 1047 <https://doi.org/10.1016/j.neuroimage.2017.04.022>
- 1048 Schneider, T., & Neumaier, A. (2001). Algorithm 808: ARFIT - A Matlab package for the estimation of parameters and  
 1049 eigenmodes of multivariate autoregressive models. *ACM Transactions on Mathematical Software*, 27(1), 58–65.  
 1050 <https://doi.org/10.1145/382043.382316>
- 1051 Scott, S. K., & Johnsrude, I. S. (2003). The neuroanatomical and functional organization of speech perception. *Trends*  
 1052 *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)
- 1053 Seth, A. K., Barrett, A. B., & Barnett, L. (2015). Granger Causality Analysis in Neuroscience and Neuroimaging.  
 1054 *Journal of Neuroscience*, 35(8), 3293–3297. <https://doi.org/10.1523/JNEUROSCI.4399-14.2015>
- 1055 Shum, M., Shiller, D. M., Baum, S. R., & Gracco, V. L. (2011). Sensorimotor integration for speech motor learning  
 1056 involves the inferior parietal cortex. *European Journal of Neuroscience*, 34(11), 1817–1822.  
 1057 <https://doi.org/10.1111/J.1460-9568.2011.07889.X>
- 1058 Skipper, J. I., Devlin, J. T., & Lametti, D. R. (2017). The hearing ear is always found close to the speaking tongue:  
 1059 Review of the role of the motor system in speech perception. *Brain and Language*, 164, 77–105.  
 1060 <https://doi.org/10.1016/J.BANDL.2016.10.004>
- 1061 Smith, A., & Weber, C. (2017). How Stuttering Develops: The Multifactorial Dynamic Pathways Theory. *Journal of*  
 1062 *Speech, Language, and Hearing Research*, 60(9), 2483–2505. [https://doi.org/10.1044/2017\\_JSLHR-S-16-0343](https://doi.org/10.1044/2017_JSLHR-S-16-0343)
- 1063 Sommer, M., Koch, M. A., Paulus, W., Weiller, C., & Büchel, C. (2002). Disconnection of speech-relevant brain areas  
 1064 in persistent developmental stuttering. *The Lancet*, 360(9330), 380–383. [https://doi.org/10.1016/S0140-](https://doi.org/10.1016/S0140-6736(02)09610-1)  
 1065 [6736\(02\)09610-1](https://doi.org/10.1016/S0140-6736(02)09610-1)
- 1066 Stasak, B., Epps, J., Schatten, H. T., Miller, I. W., Provost, E. M., & Armev, M. F. (2021). Read speech voice quality  
 1067 and disfluency in individuals with recent suicidal ideation or suicide attempt. *Speech Communication*, 132, 10–20.  
 1068 <https://doi.org/10.1016/J.SPECOM.2021.05.004>
- 1069 Strauß, A., & Schwartz, J. L. (2017). The syllable in the light of motor skills and neural oscillations. *Language,*  
 1070 *Cognition and Neuroscience*, 32(5), 562–569. <https://doi.org/10.1080/23273798.2016.1253852>
- 1071 Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production.  
 1072 *Language and Cognitive Processes*, 26(7), 952–981. <https://doi.org/10.1080/01690960903498424>
- 1073 Turk, A. Z., Lotfi Marchoubeh, M., Fritsch, I., Maguire, G. A., & SheikhBahaei, S. (2021). Dopamine, vocalization,  
 1074 and astrocytes. *Brain and Language*, 219, 104970. <https://doi.org/10.1016/j.bandl.2021.104970>
- 1075 Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M.  
 1076 (2002). Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of  
 1077 the MNI MRI Single-Subject Brain. *NeuroImage*, 15(1), 273–289. <https://doi.org/10.1006/NIMG.2001.0978>
- 1078 Watkins, K. E., Smith, S. M., Davis, S., & Howell, P. (2008). Structural and functional abnormalities of the motor  
 1079 system in developmental stuttering. *Brain*, 131(1), 50–59. <https://doi.org/10.1093/BRAIN/AWM241>
- 1080 Weiss, A. L. (1995). Conversational demands and their effects on fluency and stuttering. *Topics in Language Disorders*,  
 1081 15(3), 18–31. <https://doi.org/10.1097/00011363-199505000-00004>
- 1082 Westner, B. U., Dalal, S. S., Gramfort, A., Litvak, V., Mosher, J. C., Oostenveld, R., & Schoffelen, J. M. (2022). A  
 1083 unified view on beamformers for M/EEG source reconstruction. *NeuroImage*, 246, 118789.  
 1084 <https://doi.org/10.1016/J.NEUROIMAGE.2021.118789>
- 1085 Yairi, E., & Ambrose, N. (2013). Epidemiology of stuttering: 21st century advances. *Journal of Fluency Disorders*,  
 1086 38(2), 66–87. <https://doi.org/10.1016/J.JFLUDIS.2012.11.002>
- 1087 Yang, Y., Jia, F., Siok, W. T., & Tan, L. H. (2016). Altered functional connectivity in persistent developmental

1088 stuttering. *Scientific Reports*, 6(1), 1–8. <https://doi.org/10.1038/srep19128>  
1089 Zhang, N., Yin, Y., Jiang, Y., & Huang, C. (2022). Reinvestigating the neural bases involved in speech  
1090 production of stutterers: An ALE meta-analysis. *Brain Sciences*, 12(8), 1030.  
1091 <https://doi.org/10.3390/brainsci12081030>

1092 **FIGURE LEGENDS**

1093

1094 **Figure 1. Experimental design.** Participants listened to sentence frames and target final words in the  
1095 listening-only task (with occasional true/false judgment questions), while they had to complete the  
1096 sentence frame by naming a picture (representing the target word) in the listening-for-speaking task.  
1097 Indexes related to cortical tracking of speech and cortical connectivity were measured during the  
1098 auditory presentation of the sentence frames. Response times at picture onset in the listening-for-  
1099 speaking task were collected.

1100

1101 **Figure 2. Regions of interest (ROIs) selected for statistical comparison between groups in the**  
1102 **source and connectivity analyses.** Five ROIs were selected in the left and right hemisphere: i) the  
1103 inferior frontal gyrus (IFG, comprising the pars opercularis, triangularis, and orbitalis), ii) the  
1104 premotor and supplementary motor cortex (preM), iii) the somato-motor strip (SM), iv) the inferior  
1105 parietal lobule (IPL), and v) the superior temporal gyrus (STG).

1106

1107 **Figure 3. Behavioral results for the listening-for-speaking task.** A) Subject-level (individual)  
1108 accuracy scores (dots) and boxplots; note: y-axis starting at 0.8. B) Single-trial response times (dots),  
1109 group-level means with error bars and density distributions.

1110

1111 **Figure 4. Spectra distribution and topographic map of the CTS at the sensor level.** A) Corrected  
1112 coherence values (coherence values converted into z-scores using the mean and standard deviation  
1113 derived from the 100 random EEG-audio combinations) in the 1 – 15 Hz frequency range can be  
1114 observed across representative sensors (C3, C5, CP3, CP5, FC3, FC5, C4, C6, CP4, CP6, FC4, FC6)  
1115 of the left (LH) and right (RH) hemisphere. B) For each group (TFA: Typical Fluent Adults; AWS:  
1116 Adults Who Stutter) and task (listening-only and listening-for-speaking), we plotted the topographic  
1117 maps of uncorrected coherence values in the theta (3 - 5 Hz) frequency band.

1118

1119 **Figure 5. Source reconstruction of the CTS values in the theta range.** For each group (TFA:  
1120 Typical Fluent Adults; AWS: Adults Who Stutter) and listening condition (listening-only and  
1121 listening-for-speaking), we plotted the source maps of CTS values (coherence values converted into  
1122 z-scores using the mean and standard deviation derived from the CTS values in all the sources) in the  
1123 theta (3 - 5 Hz) frequency band.

1124

1125 **Figure 6. Mean CTS in the 3-5 Hz range (theta band) in each region of interest for each group.**  
1126 Boxplots are overlaid with individual data points and density distributions. Each dot represents data  
1127 from one of the participants (blue for TFA and red for AWS). Boxes cover the 25th to 75th percentile  
1128 (inter-quartile range; IQR). The middle of the box represents the median. Whiskers extend from the  
1129 25th percentile and 75th percentile to cover all data points lying within 1.5 times the IQR (from the  
1130 25th and 75th percentile, respectively). Regions showing a significant group effect are marked with  
1131 an asterisk.

1132

1133 **Figure 7. Causal functional connectivity analysis.** For each task, we included the connections that  
1134 exhibited statistically significantly higher PDC values for TFA compared to AWS. We included a  
1135 seed for each of the regions of interest (IFG: inferior frontal gyrus; preM: premotor/supplementary  
1136 motor; SM: sensorimotor; IPL: inferior parietal lobule; STG: superior temporal) in both the left (LH)  
1137 and right (RH) hemisphere.

1138

1139 **Figure 8. Connectivity-RTs correlation (listening-for-speaking).** Scatterplot showing the  
1140 correlation between response times (RTs) and connectivity from the right superior temporal gyrus to  
1141 the left sensorimotor cortex.

1142 **TABLES**

1143

1144 **Table 1. Model summaries for accuracy and response times for naming in the listening-for-**  
 1145 **speaking task.**

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

1146

1147

1148 **SUPPLEMENTARY MATERIAL: Supplementary figures (Figure S1-3) and tables**  
 1149 **(Table S1-2)**