Reactive Inhibitory Control Precedes Stuttering Events

Abbreviated Title: Stuttering and Reactive Inhibitory Control

Joan Orpella[1*], Graham Flick[1], M. Florencia Assaneo[2], Liina Pylkkänen[134], David Poeppel[156], Eric S. Jackson[7*]

[1]Department of Psychology, New York University, USA; [2]Institute of Neurobiology, National Autonomous University of Mexico, Juriquilla, Querétaro, Mexico; [3]Department of Linguistics, New York University, USA; [4]NYUAD Institute, New York University Abu Dhabi, United Arab Emirates; [5]Center for Language, Music and Emotion (CLaME), New York University, New York, NY, USA; [6]Ernst Strüngmann Institute (ESI) for Neuroscience, Frankfurt, Germany; [7]Department of Communicative Sciences and Disorders, New York University, USA

*Corresponding authors

Joan Orpella

jo1358@nyu.edu

Eric S. Jackson

eric.s.jackson@nyu.edu

Author contributions: Joan Orpella: Conceptualization: Lead; Formal Analysis: Lead; Investigation: Lead; Methodology: Lead; Writing: Lead. Graham Flick: Conceptualization: Equal; Formal Analysis: Supporting; Investigation: Equal; Methodology: Supporting; Writing: Supporting. M. Florencia Assaneo: Conceptualization: Equal; Formal Analysis: Supporting; Methodology: Supporting; Writing: Supporting. Liina Pylkkänen: Conceptualization: Supporting; Writing: Supporting. David Poeppel: Conceptualization: Supporting; Writing: Supporting. Eric S. Jackson: Conceptualization: Lead; Formal Analysis: Lead; Funding Acquisition: Lead; Investigation: Lead; Methodology: Lead; Project Administration: Lead; Writing: Lead.

Competing Interests: The authors declare no competing financial interests.

Abstract

Research points to neurofunctional differences underlying fluent speech production in stutterers compared to non-stutterers. Considerably less work has focused on processes that underlie stuttered vs. fluent speech within stutterers, partly due to difficulties associated with eliciting stuttering during neuroimaging. We used magnetoencephalography to test the hypothesis that a reactive inhibitory control response in the action-stopping network precedes stuttering events. We leveraged a novel clinical interview to develop participant-specific stimuli (words identified as likely to be stuttered) known to elicit a comparable amount of stuttered and fluent trials across participants. Thirty-one adult stutterers participated (8 female). Participants completed a delayed-response task which included a cue (prior to the go cue) signaling the imminent requirement to produce the word. In response to this cue, stuttered (vs. fluent) productions resulted in greater beta power in the right presupplementary motor area (R-preSMA), a key node in the action-stopping network, a signature of reactive motor inhibition. Beta power was related to stuttering severity and was predictive of stuttering on a trial-to-trial basis. Stuttered words were associated with delayed speech initiation, also consistent with the reactive inhibition hypothesis. Post-hoc analyses revealed that words associated with stronger anticipation (independently-generated vs. researcher-assisted words words) were associated with more stuttering and greater beta power. This points to a critical relationship between reactive inhibition and stuttering anticipation such that words identified more readily by participants as likely to be stuttered elicit greater inhibition. This is the largest and most balanced investigation of stuttered and fluent speech to date. It advances understanding of the neural bases of stuttering events, and offers a target region for future studies that test the efficacy of neuromodulation on stuttering.

Keywords: stuttering; fluency; reactive inhibitory control; preSMA; MEG.

Abbreviations: dSPM = dynamic statistical parameter mapping; MEG = magnetoencephalography; R-DLPFC = right dorsolateral prefrontal cortex; SSI-4 = stuttering severity index - 4th edition; SLP = speech-language pathologist; R-preSMA = right presupplementary motor area; tDCS = transcranial direct current stimulation.

Significance Statement

The neural bases of stuttered speech are unknown, largely due to difficulties in eliciting stuttered speech reliably in the lab. We used a novel method to elicit stuttered speech during MEG recordings in adult stutterers to test the hypothesis that reactive inhibition precedes stuttered speech. We observed a neural response characteristic of reactive motor inhibition (increased beta power in the right presupplementary motor area) in response to a cue signaling the imminent requirement to produce a word likely to be stuttered. Speech initiation was also delayed for stuttered trials. Beta power was related to stuttering severity and percentage of trials stuttered, and predicted whether individual trials would be stuttered. These results contribute to a brain-based account of stuttering events and have important clinical implications, e.g., providing a target region for future neuromodulation studies.

Introduction

Most neurofunctional investigations of stuttering have assessed *trait* characteristics— differences between the fluent speech of stutterers compared to that of non-stutterers—generally revealing atypical activation in left and right perisylvian speech areas (Braun et al., 1997; De Nil et al., 2000; Kell et al., 2009; Neumann et al., 2004; Preibisch et al., 2003; Watkins et al., 2008). Fewer studies have assessed *state* characteristics, or differences associated with stuttered versus fluent speech in stutterers (Belyk et al., 2015; Budde et al., 2014). This is due in large part to the difficulty associated with eliciting stuttered speech reliably in the unnatural environments of neuroimaging experiments. The few investigations that have examined state characteristics have yielded inconsistent results, primarily due to small sample sizes and limited stuttering (see Table 1). Thus, the neural substrates of *stuttered* speech are unknown.

We recently introduced a method for eliciting stuttered speech during neuroimaging (Goldfarb et al., in press; Jackson et al., 2020). Using this approach, we showed that adult stutterers exhibit elevated activation in the right dorsolateral prefrontal cortex [R-DLPFC] prior to producing anticipated words, or words identified by participants as likely to be stuttered (Jackson et al., 2022). In line with Neef et al. (2018), we interpreted this result as a form of *proactive* inhibitory

control in response to stuttering anticipation (Jackson et al., 2022). Proactive inhibitory control refers to the ability to prevent or delay undesired actions and has been associated with activation in the R-DLPFC (Aron, 2011; Jahanshahi et al., 2015). For stuttering, the assumption is that stuttered speech is an undesired action. As a result, the speaker delays or inhibits speech initiation through proactive inhibitory control because they do not want to overtly stutter (Jackson et al., 2015, 2019, 2022; Neef et al., 2018). Delaying speech initiation affords the possibility to change course of action, for example, by stalling, substituting a word, or using a speaking strategy (Jackson et al., 2015, 2019). Many stutterers report anticipating stuttering when they have to say their name. In this situation, the speaker may implement proactive control in the seconds leading up to when they actually have to say their name to avoid potential negative listener reactions.

It has also been proposed that *reactive* inhibitory control may play a role in stuttering (Hannah & Aron, 2021; Neef et al., 2016, 2018). In contrast to proactive inhibitory control, reactive inhibitory control is an automatic and fast response to stop or delay a planned action triggered by exogenous cues (Aron, 2011; Wessel & Aron, 2017). The most reliable neural signature of this inhibitory response is elevated beta power in nodes of the action-stopping network (the right pre-supplementary motor area [R-preSMA], right inferior frontal gyrus [R-IFG], and subthalamic nucleus) in response to no-go cues or stop signals (Aron, 2011; Wessel & Aron, 2017). Following the example above, reactive inhibition may be triggered when the speaker is cued to the immediacy of having to say their name. This may happen, for instance, when the communicative partner extends their hand and says their name because, at this communicative juncture, the speaker knows that they will have to produce a word that will likely be stuttered.

In the current study, we focused on reactive inhibition. We hypothesized that reactive inhibition would be triggered when an exogenous cue signals the impending requirement to produce a word likely to be stuttered. As with proactive inhibition, the assumption is that stuttering is an undesired outcome and so the speaker reacts to the impending requirement to speak by inhibiting planned speech because they do not want to overtly stutter. This could result in delayed speech initiation time. We employed our recently introduced method to elicit a balanced amount of stuttered and fluent speech across participants (Goldfarb et al., in press; Jackson et al., 2020). In

an initial visit, participant-specific anticipated words (i.e., words identified by participants as likely to be stuttered) were determined during a clinical interview procedure. In a second visit, participants completed a delayed-response task in which they produced words likely to be stuttered while neural activity was recorded with magnetoencephalography (MEG). Critically, the go signal in this task was preceded by a cue that informed participants of the imminent requirement to speak. We hypothesized that reactive inhibitory responses would be triggered by this cue prior to overt stuttering events. We assumed that stuttered words were anticipated given stutterers' ability to predict overt stuttering events with high accuracy. Reactive inhibition would be reflected as increased beta power in cortical nodes of the action-stopping network following the cue for stuttered vs. fluent productions. The presence of both independently-generated words (i.e., words independently identified by participants as likely to be stuttered) and researcher-assisted words (i.e., words identified by participants as anticipated with researcher assistance) allowed us to probe a posteriori the relationship between anticipation and reactive inhibitory control. Assuming that independently-generated words are associated with greater anticipation than researcher-assisted words led to the secondary hypothesis that independently-generated anticipated words are related to higher levels of reactive inhibitory control than researcher-assisted anticipated words.

Materials and methods

This study was approved by the Institutional Review Board at New York University. Written consent was obtained from all participants in accordance with the Declaration of Helsinki.

Participants

Participants were recruited via the last author's database of stutterers, a mass email from the National Stuttering Association, and word of mouth. Participants included 31 adults who stutter (8 female), with a mean age of 30.1 (SD = 7.8), who all reported a negative history of concomitant neurological, psychiatric, or psychological disorders. All participants reported that they were native English speakers. Stuttering diagnosis was made by the last author, a speech-language pathologist (SLP) with 14 years of experience and expertise in stuttering intervention. All participants also self-reported as a stutterer and exhibited three or more stuttering-like

disfluencies[20] with temporally aligned physical concomitants (e.g., eye blinking, head movements) during a 5-10 minute conversation. The Stuttering Severity Index - 4th Edition (SSI-4)(Riley, 2009) was also administered, and all participants responded to three subjective severity questions via 5-pt. Likert scales: 1) How severe would you rate your stuttering?; 2) How severe would other people rate your stuttering?; 3) Overall, how much does stuttering impact your life? (1 = Mild, 5 = Severe). There were two visits. Visit 1 included diagnostic testing and a clinical interview to determine participant-specific stimuli, i.e., anticipated words – words likely to be stuttered – so that the likelihood of stuttering during MEG testing was increased. Visit 2 included MEG testing.

Clinical Interview

The interview was adapted from Jackson et al. (2020) In that study, both anticipated and unanticipated words were elicited from participants, which yielded a near equal distribution of stuttered and fluent speech during fNIRS recording (Jackson et al., 2022). However, Jackson et al. (2020, 2022) included interactive speech whereas the current study does not; the likelihood of stuttering during testing with face-to-face communication is higher. Therefore, we only elicited anticipated words in the current study to increase the probability of a near-balanced distribution of stuttered and fluent speech in the absence of face-to-face communication (i.e., in the shielded room while MEG data were recorded). The interview is fully described in Jackson et al. (2020), but is also summarized here. Participants were initially asked if they anticipated stuttering; all participants confirmed that they did. Participants were then asked to identify words that they anticipate stuttering on. Most participants identified at least a few words, though there was variability across participants (as in Jackson et al.) Participants were also provided with prompts (e.g., “What about your name?”, “Is your job title a difficult word for you?”). Participants were then asked about anticipated sounds, i.e., word-initial sounds that are problematic, which were used to create additional anticipated words beginning with these sounds. This was done ultimately to produce a list of 50 different, participant-specific words to be presented during MEG testing (visit 2). Critically, all 50 words were anticipated, regardless of how they were elicited.

Stimuli

Each participant had their own list of 50 anticipated words, which were presented during the MEG experiment. While all of these words were considered anticipated words, there were words that were identified by the participant independently (i.e., *independently-generated* words), and words that were generated with researcher assistance (i.e., *researcher-assisted* words). For example, researcher-assisted words could be identified by asking participants whether specific words are difficult or asking participants about difficult sounds, and then identifying words that began with these difficult sounds. Researcher-assisted words were developed using an online word generator and were five syllables in length, because longer words tend to be stuttered more than shorter words (Brown & Moren, 1942). For example, if /b/ was identified as an anticipated sound, words like *biochemistry*, *biological*, and *biographical* may have been included. Independently-generated words were typically shorter than researcher-assisted words. Independently-generated and researcher-assisted anticipated words were each produced six times during the task.

Task

The behavioral task is depicted in Fig. 1. Each trial began with a fixation cross (baseline period) of variable duration (1 – 1.5 sec). A word from the anticipated words list (see Stimuli section) appeared in the center of the screen (0.4 sec) followed by a blank screen of variable duration (0.4 – 0.8 sec). After this blank screen, there was either a *speak* trial or a *catch* trial. For speak trials, a white asterisk appeared (henceforth, cue), signaling the requirement to speak the word at the onset of the following green asterisk (henceforth, go-signal). The duration of the white asterisk was 0.2 seconds and was always followed by 0.8 seconds of blank screen. The time between the onsets of the white asterisk (cue) and the green asterisk (go signal) was therefore always 1 second. The duration of the green asterisk on the screen was 0.5 seconds and was followed by a variable blank period (2 – 3 sec) to allow the participant to produce the word. The word STOP appeared at the end of this blank period, requiring participants to abort any incomplete speech acts and prepare for the next trial by remaining as still as possible. Catch trials (15% of trials) were introduced as a means to create uncertainty about the requirement to speak the anticipated word. In these catch trials, a red asterisk followed the blank screen and the participant was required to remain silent and await the following trial (Fig. 1). The overall design thus simulates

a common experience of people who stutter–the stutterer is required to produce an anticipated word (e.g., one's own name) within an expected time window (e.g., after the speaking partner extends their hand and says "What is your name?"). The critical window for analysis was the cue period of 1 sec before the go signal (i.e., between the white and the green asterisks). The task consisted of six blocks; each block included the list of 50 words presented in randomized order, for a total of 300 words per MEG session. Participants' faces and speech during the experiment were recorded.

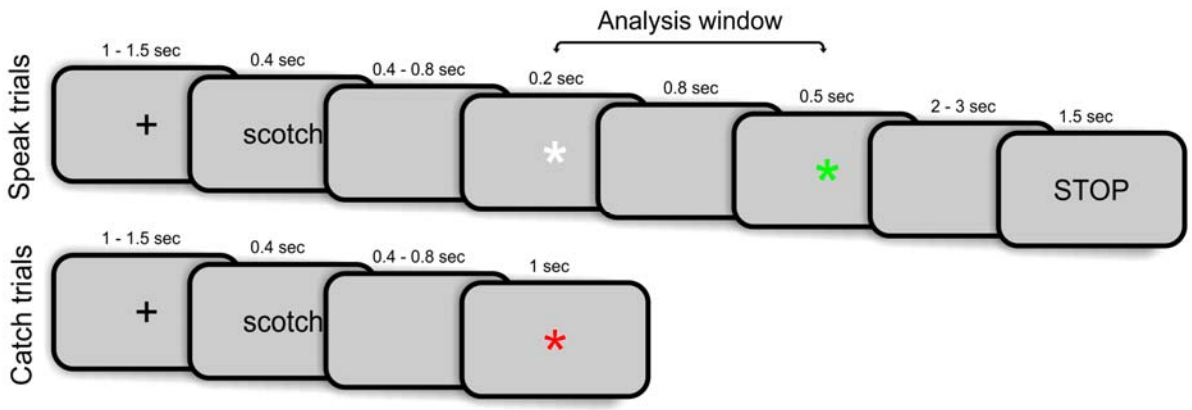

Fig 1. Behavioral task. Each trial began with a fixation cross of variable duration (Baseline period). Stimulus words appeared in the center of the screen followed by a blank screen of variable duration. For speak trials, a white asterisk appeared (cue), signaling the requirement to speak the word on the following green asterisk (go signal). Participants had 2 – 3 s to produce the words. Catch trials started in the same manner, however, a red asterisk appeared after the initial blank screen, indicating that participants should remain silent until the next trial.

MEG data acquisition and preprocessing

Neuromagnetic responses were acquired using a 157-channel whole-head axial gradiometer (Kanazawa Institute of Technology, Japan) situated in a magnetically shielded room, with a sampling rate of 1000Hz. To monitor head position during the recordings, five electromagnetic coils were attached to the subject's head. We registered the location of these coils with respect to the MEG sensors before and after each block of the experiment. Participants' head shape was digitized immediately before the MEG recordings using a Polhemus digitizer and 3D digitizer software (Source Signal Imaging) along with 5 fiducial points, to align the position of the coils with participants' head shape, and 3 anatomical landmarks (nasion, and bilateral tragus), to

further allow for the co-registration of participant's MEG data with an anatomical MRI template. An online band-pass filter (1Hz-200Hz) was applied to all MEG recordings.

Data preprocessing was conducted using custom Python scripts and MNE-python software (Gramfort et al., 2014). Bad channels were first selected and interpolated using spherical spline interpolation. A least-squares projection was then fitted to the data from a 2-minute empty room recording acquired at the beginning of each MEG session and the corresponding component was removed. MEG signals were next digitally low-pass filtered at 50Hz using MNE-python's default parameters with firwin design and finally epoched between -1700 ms and 1500 ms relative to the onset of presentation of the cue (white asterisk; Fig. 1). Linear detrending was applied to the epochs to account for signal drift. Baseline correction was applied at the analysis phase (see below). An independent component analysis was used to correct for cardiac, ocular, and muscle artifacts. The epochs resulting from these steps were visually inspected and remaining artifactual trials were discarded from further analysis.

Data Analysis - Behavioral

Trials were judged to be stuttered, fluent, or errors by the last author, an SLP with 14 years of experience and expertise in stuttering intervention. Thirty percent of the trials (randomly selected) were also classified by another SLP, blind to the study, to determine interrater reliability. Stuttered trials were those with stuttering-like disfluencies including blocks, prolongations, or part-word repetitions. Error trials were those in which participants forgot or did not attempt to produce the target word. Reaction time was calculated as the time between the go signal (i.e., the green asterisk) and speech onset as defined by the first articulatory movement or accessory behavior (as in Jackson et al., 2022)). Accessory behaviors refer to non-speech behaviors that often co-occur with stuttering events (e.g., eye blinking, facial tension). Articulatory onset was defined as the first articulatory movement, which was determined based on visual inspection using Davinci Resolve (Black Magic Design, Australia). This allowed for frame-by-frame scanning (29.97 frames per second) of the recordings of participants' faces. A generalized linear mixed model fit by maximum likelihood (family = binomial) in R(R Core Team, 2017) was used to assess variables that contributed to stuttered speech. Fixed factors included stimulus type (independently-generated or researcher-assisted), word length (number of

letters), initial phoneme (consonant or vowel), and trial number, while participant was a random factor and random slopes for stimulus by participant were also included. Multiple linear regression was used to assess whether SSI-4 score, the number of independently generated anticipated words, and stuttering impact (Q3) contribute to the percentage of trials stuttered for each participant. Finally, a linear mixed model was used to assess potential factors impacting reaction time. Speech outcome was a fixed factor and participant was a random factor.

Data Analysis - MEG

Time-frequency analysis in sensor space

To determine differences in beta power between stuttered and fluent trials, we conducted a time-frequency analysis. Prior to the decomposition, trial types were equalized in counts using MNE-python's equalize_epoch_counts function to avoid biases in the time-frequency decomposition toward either condition (*stuttered*, *fluent*). In brief, an equal number of trials from each condition is selected while minimizing the differences in the event times of the two conditions to prevent differences due to time-varying characteristics in the data (e.g., adaptation, noise). For the decomposition, we used a Stockwell transform (Stockwell, 2007) with a Gaussian window of width = 0.5, which offers a good balance between temporal and frequency resolution. The decomposition was performed for each condition separately (*stuttered*, *fluent*), for frequencies between 12 and 30 Hz (beta frequency band) and times between the cue presentation and the presentation of the go signal (Fig. 1). A baseline correction was applied at this stage by subtracting the mean of the values between -1.6 and -1.3 prior to the cue presentation (i.e., within the baseline period; Fig. 1) followed by dividing by the mean of the same values (i.e., yielding percent change from baseline). Note that due to the jitter following word presentation, the baseline period differed for each trial. The contrast between *stuttered* and *fluent* trials was performed by direct subtraction (*stuttered* minus *fluent*) within participant before averaging the resulting difference across participants. To determine the statistical significance of observed differences, we first entered the difference data for each participant into a cluster-based permutation test (two-tailed; 1,024 permutations) across participants using MNE-python's spatio-temporal_cluster_1samp_test function, with an initial t-threshold of 2.05 (equivalent to $p < 0.05$ for the number of participants) and a subsequent cluster threshold of $p < 0.05$. The same

analysis was performed as a control to determine differences in the theta (4 - 8 Hz), alpha (9 - 12 Hz), and low gamma (30 - 50 Hz) bands during the same time window (cue presentation to go signal).

Correlation between overall beta power and behavioral measures

Spearman's Rank correlation coefficient was used to evaluate the relationship between the normalized beta power difference between stuttered and fluent trials across sensors and stuttering severity. Specifically, we used participants' beta power difference between stuttered and fluent trials within 0.2 and 0.5 seconds and 22 to 27 Hz and their SSI-4 score (Table 2). One participant had missing data for the SSI-4, and was excluded from this analysis. Statistical significance was established by a threshold of $p < 0.05$. We identified an outlier in the data by a criterion of 2 SD from the mean of the beta power differences. Results in-text are reported without the outlier and both with and without the outlier in the corresponding figure. Correlations were also evaluated between participants' SSI-4 score and beta power relative to baseline in stuttered and fluent trials separately. In addition, we assessed the correlation between beta power difference and participants' number of independently-generated words. Statistical significance was established by a threshold of $p < 0.05$ in all cases.

Power-spectral density in source space

To determine the cortical origins of observed power differences between *stuttered* and *fluent* trials, we projected each participant's epoched data to an *fsaverage* source space template (ICO 4). For each participant, we used the standard MNE-python pipeline to compute a forward model based on a 1-layer boundary element model and a minimum-norm inverse model (signal-to-noise ratio = 2; loose dipole fitting = 0.2, normal orientation) using a noise covariance matrix computed from all sensors averaged over a baseline period of 300 ms across trials and determined based on a log-likelihood and cross-validation procedure(Engemann & Gramfort, 2015). This inverse model was applied to the participant's epochs (*stuttered* and *fluent* separately, equalized in counts) using dynamic statistical parameter mapping (dSPM). Power spectral density in source space was estimated for each trial of each condition and using a multi-taper method. A slightly larger time-window than the time-frequency analysis (0.1 to 0.5 seconds after the cue) was used for a better estimation of power; the frequency range was kept to 22 to 27

Hz. We computed the average across trials for each condition from which the normalized difference in power [(*stuttered – fluent)* / (*stuttered* + *fluent*)] was estimated. Each participant's data was then morphed to a common template and entered into a cluster-based permutation test (one-tailed; 1,024 permutations) across participants (N = 29) with an initial *t*-threshold equivalent to $p < 0.01$ and a subsequent cluster threshold of $p < 0.05$.

Logistic regression analysis

We assessed the extent to which the trial-by-trial modulations of beta power in the R-preSMA, compared to baseline, predicted the corresponding trial's outcome (stuttered or fluent). For each participant and trial, we computed the normalized difference between the beta power (22 to 27 Hz) within R-preSMA in the identified times during the cue period (0.1 to 0.5 seconds) and in an equivalent 400ms period within the baseline (-1.6 to -1.2 seconds before the cue). These trial-wise power difference values were used to predict the outcome (stuttered or fluent) of the corresponding trials using logistic regression. Specifically, we fit and scored a logistic regression model per participant. Scores for each participant's model were determined based on the number of correct predictions, such that if *j* outcomes are predicted correctly out of n observations (x values) the model's score is j/n. Participants' scores were then entered into a one-sample *t*-test against chance (0.5) to determine statistical significance across participants.

Correlation between R-preSMA beta power and percent trials stuttered

Spearman's Rank correlation coefficient was used to evaluate the relationship between beta power with respect to baseline in the identified pre-SMA cluster and percent trials stuttered. To better characterize the relationship between these variables, we fit both a linear and an exponential regression model. Two outliers were identified as falling beyond the regression model's prediction interval. The R^2 of both linear and exponential models is reported with and without the effect of the outliers.

Whole-brain event-related fields analysis

In addition to the hypothesis-driven analyses in frequency space, we explored the differences in event related fields (ERFs) associated with stuttered and fluent trials using the broadband signal. MEG signals were digitally low-pass filtered at 30Hz using MNE-python's default parameters with firwin design and epoched between -1700 ms and 1500 ms relative to the onset of

presentation of the cue (white asterisk; Fig. 1). Epochs were mean-centered, linearly detrended, and equalized in counts as above so that each participant had the same number of epochs per trial type (stuttered, fluent). We applied each participant's previously computed inverse model (see *Power-spectral density in source space* section above) to the average across trials of each condition (i.e., its ERF) and each projection was morphed to a common (fsaverage) space. Finally, we computed the normalized difference between the trial types as (stuttered-fluent)/(stuttered+fluent) before averaging across subjects. To explore differences between the trial types between the cue presentation and the go signal, we performed a series of one-sample cluster-based permutation tests (1000 permutations; two-tailed) across participants (N = 29) with an initial *t*-threshold equivalent to $p < 0.01$ and a subsequent cluster threshold of $p < 0.05$ as before. Specifically, the tests were performed at the whole-brain level (i.e., using all 5124 sources) using a sliding window of 100 ms with a temporal overlap of 50 ms (i.e., sliding forward in increments of 50 ms). Results are reported for clusters that exceeded the $p < 0.05$ threshold.

Data availability

All data will be made available upon the paper's acceptance.

Results

Behavioral

Out of 7,964 trials, 3,166 trials were stuttered (39.75%) (across participants, SD = 24.01%; MIN = 3%; MAX = 92%), and 4,798 trials (60.25%) were produced fluently. 54.68% of independently-generated anticipated words were stuttered (SD = 37.07%; MIN = 0%; MAX = 100%), whereas 37.71% of researcher-assisted anticipated words were stuttered (SD = 24.28%; MIN = 3%; MAX = 92%). Interrater reliability for 30% of the trials, between the final author and a SLP blind to the study, yielded a Cohen's weighted kappa of .93 ($p < .05$), indicating strong agreement (McHugh, 2012). Mean reaction time was 675.84 ms (SD = 272.99). Interrater reliability for reaction time was tested by randomly selecting 30% of the trials and calculating

Cohen's weighted kappa (k = 0.85, p < .05), indicating strong agreement (McHugh, 2012). Due to the challenges associated with identifying anticipated words, there was significant variability between participants in the number of independently-generated anticipated words (M = 4.90; SD = 7.96). More stuttering was elicited for independently-generated vs. researcher-assisted words ($\boldsymbol{\beta}$ = 0.86, z = 6.57, p < .001). Both word length and initial phoneme contributed to stuttered vs. fluent trials ($\boldsymbol{\beta}$ = -0.03, z = -2.72, p < .001; $\boldsymbol{\beta}$ = -0.32, z = -4.07, p < .001, respectively) such that longer words were stuttered *less frequently*, and words starting with consonants were stuttered more frequently than words starting with vowels. The word length finding is not in line with the literature (Brown & Moren, 1942), but this may be because independently-generated words were generally shorter than researcher-assisted words. The initial phoneme result is in line with the literature (Sheehan, 1974; Taylor, 1966). Overt stuttering severity (SSI-4 score) predicted percent trials stuttered (□ = 1.72, t = 3.36, p = .003), whereas the number of independently-generated anticipated words and stuttering impact (Q3) did not (□ = -0.10, t = -0.18, p = 0.86, □ = 0.04, t = 0.01, p = .99, respectively). Finally, stuttered trials were associated with greater reaction time (or speech initiation time) than fluent trials (□ = 0.82, t = 2.25, p < .05). Table 2 includes total percent trials stuttered, percent trials stuttered for independently-generated and researcher-assisted anticipated words, SSI-4 total scores and rating, responses to each subjective severity question (Q1-Q3), and reaction time.

Neural

Time-frequency in sensor space

Two participants did not attend their scheduled MEG visits (i.e., they only attended the first visit with clinical interview and diagnostic testing), thus all MEG analyses include 29 participants. A cluster-based permutation test (N = 29, 1,024 permutations; two-tailed) on the difference between stuttered and fluent trials (*Materials and methods*) indicated a significant difference in beta power (p < 0.05) with greater power in stuttered trials in the range between 0.2 and 0.5 seconds after the cue and between 22 and 27 Hz (Fig. 2A). This effect remained significant (p < 0.05; one-tailed) after removing an outlier that was 2 SD above the mean beta power difference in the identified cluster (*see Materials and Methods*). No significant differences in power were found in the theta (4 - 8 Hz), alpha (9 - 12 Hz), or low gamma (30 - 50 Hz) bands. We assessed

the relationship between beta power differences within the reported cluster (22 to 27 Hz and 0.2 to 0.5 seconds) and stuttering severity (indexed by SSI-4 scores). A significant correlation was found between beta power differences and stuttering severity (Spearman rho = 0.38, $p < 0.05$; Fig. 2B). No significant correlations were found between severity scores and beta power in either stuttered or fluent trials independently.

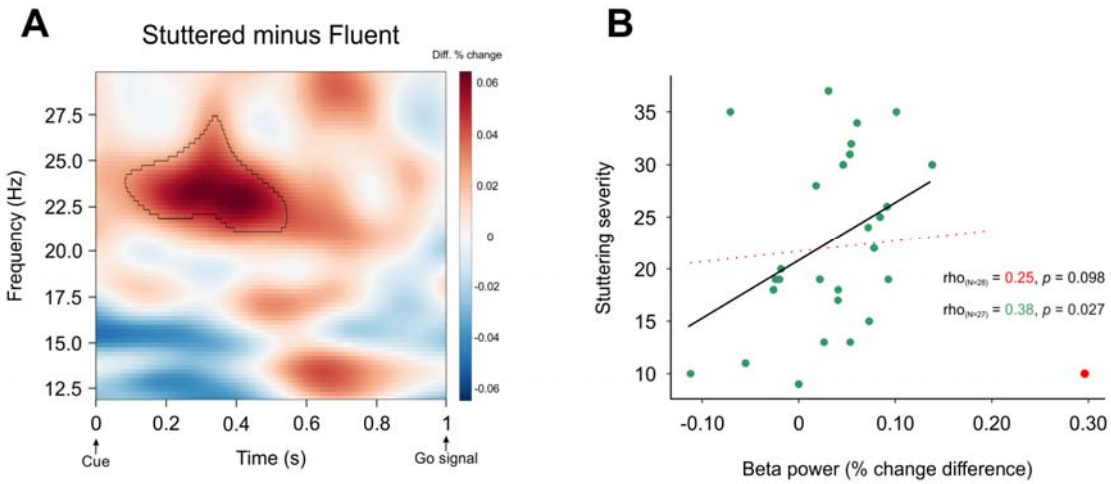

Fig 2. Time frequency analysis. A. Greater beta power prior to stuttered compared to fluent speech. The plot shows the difference (*stuttered* minus *fluent*) time-frequency plot with the significant cluster in the beta band highlighted (black contour). Significance was determined via a cluster-based permutation test ($p < 0.05$; two-tailed; 1,024 permutations) across participants (N = 29). The time-frequency analysis window spanned 1 second between the cue (white asterisk; Fig. 1) and the go signal (green asterisk). Zero (0) is the time when the cue was presented while 1 is the time when the go-signal was presented. **B. Correlation between beta power and stuttering severity.** Scatterplot of the relationship between SSI-4 scores (a proxy for stuttering severity) and beta power (22 to 27 Hz, 0.2 to 0.5 seconds after the cue). Spearman rho values and corresponding *p* values reported in the plot for the correlation without (green) and with (red) the outlier (red dot at lower right).

There was also a significant correlation between beta power differences and participants' number of independently-generated words (Spearman rho = 0.34, $p < 0.05$).

Power-spectral density in source space

To determine the cortical origin of beta power differences between stuttered and fluent trials, we computed the normalized difference in power [(*stuttered – fluent)* / (*stuttered* + *fluent*)] from each participant's average power spectral density for each condition in source space (*Materials*

and methods). A cluster-based permutation test (N = 29; 1,024 permutations; one-tailed) across participants using participants' source projected data (the entire cortical surface morphed to a common *fsaverage* space) spanning 0.1 to 0.5 seconds after the cue and a 22 - 27 Hz frequency range (*Materials and methods)* yielded significant differences ($p < 0.05$) in a single cluster comprising the right anterior and medial SMA (i.e., preSMA; Fig. 3). The precise location of the cluster in cortical space was determined by quantifying the number of sources in the cluster falling within each label of the Glasser et al.'s (2016) cortical atlas. The largest number of sources fell into label 6ma (medial anterior) in the right hemisphere, corresponding to the R-preSMA. Zero sources fell into the right SMA proper (label 6mp [medial posterior]).

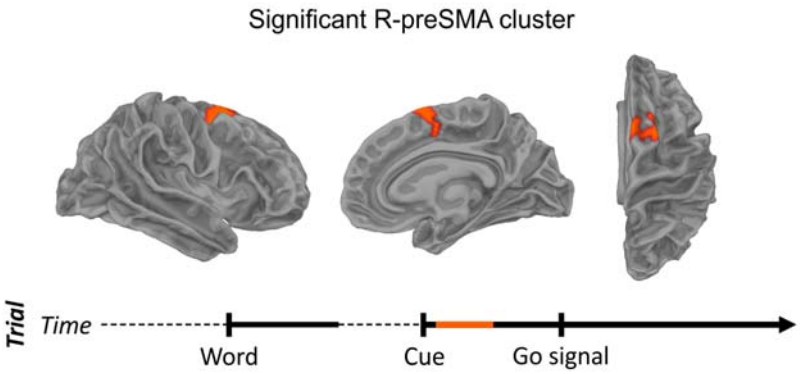

Fig 3. Greater beta power in the R-preSMA prior to stuttered speech. Significant cluster shown in a lateral (**left**), medial (**middle**), and axial (**right**) views. Significance was determined via a cluster-based permutation test across participants (one-tail) using each participant's power spectral density estimates in source space (5124 sources) averaged over frequencies 22 – 27 Hz and times 0.1 to 0.5 sec after the cue, with a *t*-threshold of $p < 0.01$ and a subsequent cluster threshold of $p < 0.05$. Timeline shows pertinent window of interest in orange.

Relationship between R-preSMA power and stuttering outcome

These results are in line with the hypothesis that stuttered speech is related to higher beta power *prior* to speech initiation in the R-preSMA, consistent with a motor inhibition response in reaction to the imminent requirement to speak an anticipated word.

Based on these results, we next evaluated the extent to which each subject's trial-by-trial beta power modulations in the R-preSMA, with respect to baseline, predicted the trial's outcome (fluent or stuttered). For each participant, we fit and scored a logistic regression model between the trial-wise beta power compared to baseline and the trials' outcome (*Materials and methods*).

A one-sample *t*-test across subjects against chance (0.5) on the scores of each subject's logistic regression model indicated that the trial-by-trial beta power modulations relative to baseline predicted whether speech would be subsequently stuttered ($t(28) = 3.529$, $p < 0.001$).

We next assessed the relationship between beta power within the R-preSMA cluster and percentage of stuttered trials. We found a positive linear correlation between each participant's mean beta power across trials (relative to baseline; *Materials and methods*) and percent trials stuttered (Spearman rho = 0.5, $p < 0.01$; Fig. 4), indicating that increased beta power in the R-preSMA is associated with more stuttering. The superior fit of an exponential regression model ($R^2 = 0.80$; Fig. 4, right) over a linear regression model ($R^2 = 0.60$; Fig. 4, left) indicates that the percentage of trials stuttered increased exponentially with beta power increases in the R-preSMA with respect to baseline.

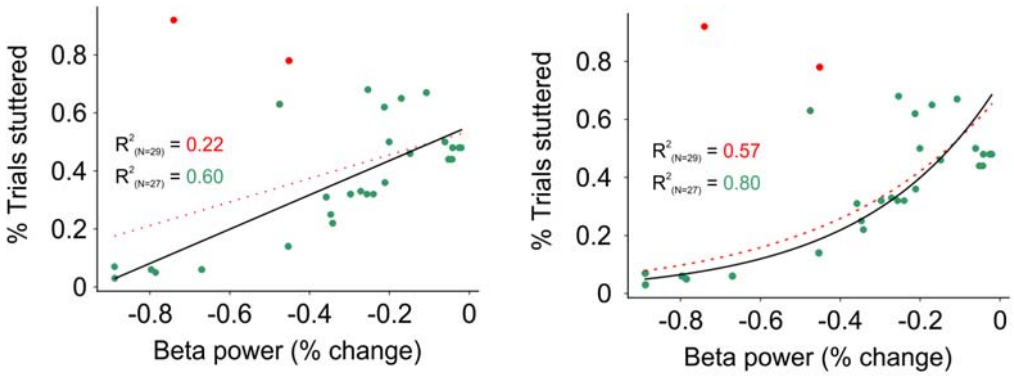

Fig 4. R-preSMA beta power predicts the percentage of trials stuttered. Left: Linear regression between beta power changes from baseline in the R-preSMA and percentage of trials stuttered. The plot shows a positive linear relationship both with (solid black line) and without (dotted red line) the 2 outliers (red dots; *Materials and methods*). **Right:** Exponential fit to the same data as the left plot, with (solid black line) and without (dotted red line) the outliers. The greater variance explained (R^2 values) for the exponential fit suggests that the percentage of trials stuttered grows exponentially with beta power increases from baseline.

In addition to the hypothesis-driven analyses discussed in this section, we conducted a whole-brain analysis of the differences between stuttered and fluent trials (see *Whole brain event-related fields analysis*; Fig. 5). A similar pre-SMA cluster emerged between 200 ms and 300 ms after the cue in line with the time-frequency and power spectral density results (Fig 5 inset). R-preSMA activity was followed by significant activity in the right dorsolateral prefrontal cortex

(R-DLPFC) 500 ms after the cue as well as activity in the left SMA and posterior superior temporal gyrus which may relate to the planning or initiation of speech.

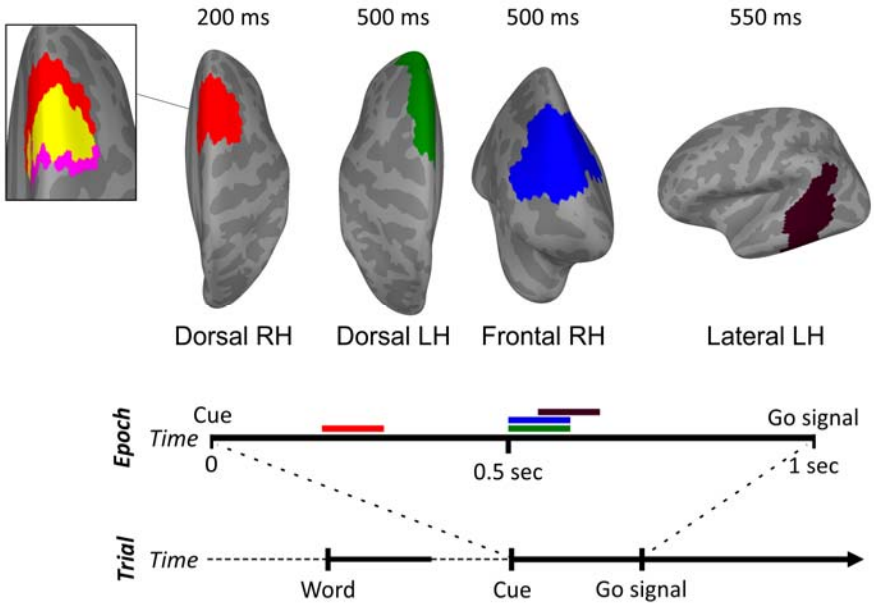

Fig 5. Whole-brain ERF analysis results. The analysis yielded four distinct clusters with greater activity in stuttered compared to fluent trials ($p < 0.05$). Between 200 and 300 ms after the Cue, there was greater activity for stuttered trials in the right dorsal frontal cortex (red cluster), in a cluster largely overlapping in time and spatial location that identified in the PSD analysis (inset: preSMA cluster from Fig. 3 [magenta]; overlap [yellow]). This was followed by greater activity for stuttered trials between 500 and 600 ms in the left superior frontal gyrus (green cluster) and in the right dorsolateral prefrontal cortex (blue cluster). There was also greater activity for stuttered trials in the left posterior superior and middle temporal cortex (dark violet cluster) between 550 and 650 ms. The timeline of events in the experimental trial and the temporal segment corresponding to the epoched data (0 to 1 second after Cue presentation) are shown below the clusters. The approximate times in which significant differences were found are indicated by horizontal bars with the color of each cluster. RH: Right Hemisphere. LH: Left Hemisphere.

Discussion

This study examined the neurofunctional substrates of stuttered speech. We tested the hypothesis that a reactive motor inhibition response consistent with action-stopping precedes stuttered speech and results from a cue signaling the impending requirement to produce a word that will subsequently be stuttered. We leveraged a recently introduced clinical interview procedure

(Goldfarb et al., in press; Jackson et al., 2020) to elicit a relatively balanced number of stuttered and fluent trials across participants, making this the largest and most balanced neurofunctional investigation of stuttered speech to date. Consistent with our hypothesis, we found that stuttered speech is preceded by increase beta power in the R-preSMA, a neural response characteristic of reactive motor inhibition. This neural response was time-locked to a cue signaling the impending requirement to produce speech. Trial-by-trial beta power in the R-preSMA predicted whether a trial was stuttered or fluent. Stuttered trials were associated with delayed speech initiation as defined by the first speech-related movement or acoustic activity. These results suggest a slowing of the motor system which may be a consequence of reactive inhibition. Secondary analyses revealed that independently-generated anticipated words were associated with greater beta power and more stuttering than researcher-assisted anticipated words pointing to a relationship between self-perceived likelihood of stuttering and reactive motor inhibition.

The primary finding of this study is that after a cue signaling the imminent requirement to speak but prior to speech execution, stuttered trials are associated with increased beta power in the R-preSMA. This finding is in line with previous suggestions that stuttering is associated with a hyperactive inhibitory control system within the cortico-basal ganglia-thalamo-cortical loop (CBGTC) which interferes with the smooth execution of successive speech movements (Hannah & Aron, 2021; Neef et al., 2016, 2018). Neef et al. (2018) found greater structural connectivity in the right frontal aslant tract in stutterers vs. non-stutterers. The frontal aslant tract connects the IFG with SMA complex which have both been associated with reactive inhibitory control (Dick et al., 2019; Wessel & Aron, 2017). While we observed greater activity in the R-preSMA, we did not find elevated activity in the R-IFG. However, the role of the R-IFG in reactive inhibition is somewhat unclear (Aron et al., 2014) which may explain this result. Korzeczek et al. (2022) recently examined the possibility of hyperactive inhibition in stuttering using a design that featured a cue prompting participants to prepare to speak (like in the current study) followed by a go-signal to speak. They did not report evidence of hyperactive motor inhibition in the contrast between stuttered and fluent speech (Korzeczek et al., 2022), but this may have been due to a limited number of stuttered trials and the use of pseudowords. Nevertheless, Korzeczek et al. reported increased beta power from midline electrodes in participants with severe vs. mild stuttering which was interpreted as a reactive inhibitory response. Korzeczek et al. did not localize the source of this activity, although they suggested a R-preSMA or R-IFG origin. Our

results extend previous accounts of hyperactive inhibition and stuttering (Hannah & Aron, 2021; Neef et al., 2016, 2018) by testing the hypothesis in the context of stuttered speech.

It is important to highlight that the cause of stuttered speech remains unknown. Proponents of the hyperactive inhibition account suggest that hyperactive inhibition may cause the interruptions in speech characterized as stuttering (i.e., blocks, prolongations, repetitions) by hindering the initiation or sequencing of speech movements (Hannah & Aron, 2021; Neef et al., 2016, 2018). The question then is what causes the inhibitory response, or why such an inhibitory response would be initiated in stuttering. In the current work, we provide an account of why and how inhibitory control may be triggered and contribute to the overt symptoms of stuttering. Stutterers recognize certain words as likely to be stuttered (i.e., anticipated words) due to a developmental process through which they learn associations between words that are stuttered and instances of stuttering (Arenas, 2017; Garcia-Barrera & Davidow, 2015). When the speaker is given a cue of the imminent requirement to produce one of these words, reactive inhibition is triggered because the speaker, instinctively, does not want to produce the word (i.e., does not want to stutter). The cortical origin (R-preSMA) and timing (~200-400 ms after the cue) of the response reported herein indicate a fast and automatic response consistent with reactive inhibition. This mirrors stop signal tasks insofar as stop signals and the cue in the current study are both associated with the desire to *not* move forward with the previously planned action. However, after this initial response, the tasks are quite different such that for the stop signal task, the task is completed while for the current task, the participant is still required to produce the word. The delayed speech initiation in the current study for stuttered vs. fluent productions may thus result from this reactive inhibitory response. However, we do not believe that reactive inhibitory control causes stuttering, but rather suggest that inhibitory control shapes the overt stuttering event, and therefore may relate to neural processes largely independent from those that *cause* the stuttering event. It may be that the cause of stuttering events relates to a dysfunction in the left hemisphere CBGTC loop for speech motor control, as proposed by Chang and Guenther (Chang & Guenther, 2020). It is possible that this CBGTC dysfunction is present near the onset of stuttering in early childhood and that hyperactive inhibitory control develops throughout childhood as a response to experiencing the intermittent interruptions characteristic of stuttering. Hyperactive inhibitory control could also interfere with speech motor control, in a way similar to Alm's (2014) proposal that social cognition disrupts an already vulnerable speech motor control system.

Neef et al. (2018) suggested that in addition to reactive control, proactive inhibitory control may be related to stuttering. Proactive control is prospective and refers to the ability to prevent or delay undesired actions (Aron, 2011; Jahanshahi et al., 2015). If a speaker knows that at some point in the future they will have to produce a word likely to be stuttered, they may implement proactive inhibitory control in order to alter planned speech (e.g., by stalling, switching a word) (Jackson et al., 2015, 2019). Jackson et al. (2022) reported elevated activation in the R-DLPFC for anticipated vs. unanticipated words and interpreted this result as a form of proactive inhibitory control in response to the upcoming requirement to produce an anticipated word. In that study, the task was designed to elicit proactive inhibitory control—there was a five-second delay between when participants first knew the word that they would have to produce and the speech initiation cue. In the current study, the task was designed to elicit a fast automatic response to a cue that indicated the imminent requirement to produce speech that would likely be stuttered. In this way, it is the nature of the response as well as how it is elicited that distinguishes between proactive and reactive control. Neurally, proactive inhibitory control is likely implemented via the indirect CBGTC loop, which includes the R-DLPFC and is characterized by a slower or more gradual response (Aron, 2011; Jahanshahi et al., 2015). In contrast, reactive inhibitory control is likely implemented via the hyperdirect CBGTC pathway, which includes the R-preSMA and is characterized by faster and automatic responses (Aron, 2011; Jahanshahi et al., 2015). Although we designed our task to investigate reactive inhibition in the current study, it is likely that both reactive and proactive inhibitory control were implemented. It is possible that proactive control was initiated when the anticipated word was presented and sustained until the word was produced. This would be in line with the pattern of activation in the R-DLPFC observed by Jackson et al. (2022). Reactive inhibition, in contrast, would have been initiated automatically in response to the cue that indicated the imminent requirement to produce the word. This automatic inhibitory response occurs because the speaker instinctively does not want to produce a word that will likely be stuttered. This interpretation is consistent with our finding that independently-generated anticipated words are related to greater beta power and more instances of stuttering than researcher-assisted words presumably because they are associated with a higher likelihood of being stuttered. Both proactive and reactive inhibitory control may contribute to delayed speech initiation as we observed. There was also some evidence in the current study that the R-DLPFC was activated prior to speech initiation

(~500 ms after the cue), which further suggests concurrent inhibitory processes. Future studies should clarify the relationship between proactive and reactive control in stuttering and the time course(s) associated with the hyperdirect and indirect pathways.

Clinical Implications

The current findings, as well as other recent work (Jackson et al., 2022), have important implications for neuromodulation as a possible complement to existing behavioral forms of stuttering therapy. Transcranial direct current stimulation (tDCS) is starting to be applied in stuttering research (Chesters et al., 2018; Garnett et al., 2019; Yada et al., 2018), albeit with modest results. For example, Garnett et al. (2019) tested the impact of anodal tDCS on overt severity in 14 adult stutterers, and while they did not find significant effects on overt stuttering severity, they found that the atypically strong association between overt severity and right thalamocortical activity was attenuated after tDCS, especially in severe stutterers. It may be that the modest effects reported to date are due to an exclusive focus on the speech network. Future neuromodulation studies can target, for example, proactive (R-DLPFC) and reactive inhibition (R-preSMA) to test whether forward-moving speech is facilitated by reducing interference from hyperactive right hemisphere areas.

Considerations

There are several points to consider when interpreting the results of this manuscript. First, we assumed that a state of anticipation was present during the experiment for stuttered trials and not fluent trials but this could not be confirmed. There is currently no way to measure anticipation reliably (Jackson et al., 2022). For example, asking participants during the experiment is unreliable and adds a considerable amount of time (see Jackson et al., 2022). However, stutterers predict stuttering with extremely high accuracy, especially when there is a delay between the point at which the speaker knows the word they are going to produce and when they are given a signal to produce the word (for a review, see Bloodstein et al., 2021). Therefore, it is likely that a state of anticipation preceded all stuttered but not fluent productions in the current study.

Second, fluctuations of beta power have been associated with various processes related to motor preparation and execution. We found no differences between stuttered and fluent speech in motor beta suppression. Differences in beta power between the conditions were restricted to the high

beta band (22 Hz – 27 Hz) and localized to the R-preSMA rather than motor cortex as would be expected for a motor-related beta. The transient nature of the response and its timing also precludes that the response in the R-preSMA is due to the lateralized readiness potential (LRP) or contingent negative variation (CNV), which are slow ramping brain potentials related to motor preparation (Hultin et al., 1996; Praamstra et al., 1996). Moreover, the LRP and CNV are localized to SMA rather than preSMA (Hultin et al., 1996; Praamstra et al., 1996). A third consideration concerns the globality of the motor inhibitory response. Indeed, reactive inhibitory control has been associated with a global motor inhibition response via excitation of the subthalamic nucleus (Wessel & Aron, 2017). Whether the observed R-preSMA activity affects global versus speech-specific motor responses in the context of stuttering remains an interesting empirical question. Future research using our paradigm can investigate whether other motor effectors are affected, for example, by assessing transcranial magnetic stimulation-evoked motor potentials associated with non-speech effectors.

Conclusion

We tested the hypothesis that stuttering events are preceded by a reactive inhibitory response to the imminent requirement to produce a word that is likely to be stuttered. Consistent with this hypothesis, we observed a neural response with the characteristics of reactive inhibitory control (increased beta power in the R-preSMA) that predicts subsequent stuttering. There is also evidence that this neural response is linked to stuttering anticipation, whereby increased self-perceived likelihood of stuttering triggers reactive inhibitory control when the speaker is faced with the imminent requirement to speak. Moreover, this reactive inhibitory response was associated with delayed speech initiation, reflecting a slowing of the motor system. This work extends previous accounts of hyperactive motor inhibition and stuttering. Methodologically, the study is the largest and most balanced comparison of stuttered and fluent speech, addressing a significant challenge in stuttering research, i.e., eliciting stuttering during neuroimaging. The findings also have potential clinical implications for neuromodulation, and open the door to new studies to further investigate the brain bases of stuttered speech.

Acknowledgements

We thank all of the stutterers who participated in this study. Without their willingness to devote their time, research on stuttering would not be possible. We also thank Haley Warner and Jake Goldfarb for assistance with coding.

Funding

Eric S. Jackson, National Institute on Deafness and Other Communication Disorders, Award ID: R21DC017821; University Research Challenge Fund, NYU. Liina Pylkkänen, NYUAD Institute, New York University Abu Dhabi, Grant G1001. David Poeppel and M. Florencia Assaneo, National Science Foundation, Award ID: Grant 2043717.

References

Alm, P. A. (2014). Stuttering in relation to anxiety, temperament, and personality: Review and analysis with focus on causality. *Journal of Fluency Disorders*, *40*, 5–21.

Arenas, R. (2017). Conceptualizing and investigating the contextual variability of stuttering: The speech and monitoring interaction (SAMI) framework. *Speech, Language and Hearing*, *20*(1), 15–28.

Aron, A. R. (2011). From reactive to proactive and selective control: Developing a richer model for stopping inappropriate responses. *Biological Psychiatry*, *69*(12), e55–e68.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: One decade on. *Trends in Cognitive Sciences*, *18*(4), 177–185.

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

Bloodstein, O., Ratner, N. B., & Brundage, S. B. (2021). *A handbook on stuttering*. Plural Publishing.

Braun, A. R., Varga, M., Stager, S., Schulz, G., Selbie, S., Maisog, J. M., Carson, R. E., & Ludlow, C. L. (1997). Altered patterns of cerebral activity during speech and language production in developmental stuttering. An H2 (15) O positron emission tomography study. *Brain*, *120*(5), 761–784.

Brown, S. F., & Moren, A. (1942). The frequency of stuttering in relation to word length during oral reading. *Journal of Speech Disorders*, *7*(2), 153–159.

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

Chang, S.-E., & Guenther, F. H. (2020). Involvement of the cortico-basal ganglia-thalamocortical loop in developmental stuttering. *Frontiers in Psychology*, *10*, 3088.

Chesters, J., Möttönen, R., & Watkins, K. E. (2018). Transcranial direct current stimulation over left inferior frontal cortex improves speech fluency in adults who stutter. *Brain*, *141*(4), 1161–1171.

De Nil, L. F., Kroll, R. M., Kapur, S., & Houle, S. (2000). A positron emission tomography study of silent and oral single word reading in stuttering and nonstuttering adults. *Journal of Speech, Language and Hearing Research*, *43*(4), 1038.

Dick, A. S., Garic, D., Graziano, P., & Tremblay, P. (2019). The frontal aslant tract (FAT) and its role in speech, language and executive function. *Cortex*, *111*, 148–163.

Engemann, D. A., & Gramfort, A. (2015). Automated model selection in covariance estimation and spatial whitening of MEG and EEG signals. *NeuroImage*, *108*, 328–342.

Garcia-Barrera, M. A., & Davidow, J. H. (2015). Anticipation in stuttering: A theoretical model of the nature of stutter prediction. *Journal of Fluency Disorders*, *44*, 1–15. https://doi.org/10.1016/j.jfludis.2015.03.002

Garnett, E. O., Chow, H. M., Choo, A. L., & Chang, S.-E. (2019). Stuttering severity modulates effects of non-invasive brain stimulation in adults who stutter. *Frontiers in Human Neuroscience*, *13*, 411.

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C. F., & Jenkinson, M. (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, *536*(7615), 171–178.

Goldfarb, J., Orpella, J., & Jackson, E. S. (in press). Eliciting Stuttering in School-age and Adolescent Stutterers in Experimental Settings. *Journal of Speech, Language & Hearing Research*.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Parkkonen, L., & Hämäläinen, M. S. (2014). MNE software for processing MEG and EEG data. *Neuroimage*, *86*, 446–460.

Hannah, R., & Aron, A. R. (2021). Towards real-world generalizability of a circuit for action-stopping. *Nature Reviews Neuroscience*, *22*(9), 538–552.

Hultin, L., Rossini, P., Romani, G. L., Högstedt, P., Tecchio, F., & Pizzella, V. (1996). Neuromagnetic localization of the late component of the contingent negative variation. *Electroencephalography and Clinical Neurophysiology*, *98*(6), 435–448.

Jackson, E. S., Dravida, S., Zhang, X., Noah, J. A., Gracco, V., & Hirsch, J. (2022). Activation in Right Dorsolateral Prefrontal Cortex Underlies Stuttering Anticipation. *Neurobiology of Language*, 1–95.

Jackson, E. S., Gracco, V., & Zebrowski, P. M. (2020). Eliciting stuttering in laboratory contexts. *Journal of Speech, Language, and Hearing Research*, *63*(1), 143–150.

Jackson, E. S., Rodgers, N. H., & Rodgers, D. B. (2019). An exploratory factor analysis of action responses to stuttering anticipation. *Journal of Fluency Disorders*, *60*, 1–10.

Jackson, E. S., Yaruss, J. S., Quesal, R. W., Terranova, V., & Whalen, D. H. (2015). Responses of adults who stutter to the anticipation of stuttering. *Journal of Fluency Disorders*, *45*, 38–51. https://doi.org/10.1016/j.jfludis.2015.05.002

Jahanshahi, M., Obeso, I., Rothwell, J. C., & Obeso, J. A. (2015). A fronto–striato–subthalamic–pallidal network for goal-directed and habitual inhibition. *Nature Reviews Neuroscience*, *16*(12), 719–732.

Kell, C. A., Neumann, K., von Kriegstein, K., Posenenske, C., von Gudenberg, A. W., Euler, H., & Giraud, A.-L. (2009). How the brain repairs stuttering. *Brain*, *132*(10), 2747–2760.

Korzeczek, A., Neef, N. E., Steinmann, I., Paulus, W., & Sommer, M. (2022). Stuttering severity relates to frontotemporal low-beta synchronization during pre-speech preparation. *Clinical Neurophysiology*, *138*, 84–96.

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276–282.

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

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

Neumann, K., Euler, H. A., Gudenberg, A. W. von, Giraud, A.-L., Lanfermann, H., Gall, V., & Preibisch, C. (2004). The nature and treatment of stuttering as revealed by fMRI: A within-and between-group comparison. *Journal of Fluency Disorders*, *28*(4), 381–410.

Praamstra, P., Stegeman, D. F., Horstink, M., & Cools, A. R. (1996). Dipole source analysis suggests selective modulation of the supplementary motor area contribution to the

readiness potential. *Electroencephalography and Clinical Neurophysiology*, *98*(6), 468–477.

Preibisch, C., Neumann, K., Raab, P., Euler, H. A., von Gudenberg, A. W., Lanfermann, H., & Giraud, A.-L. (2003). Evidence for compensation for stuttering by the right frontal operculum. *Neuroimage*, *20*(2), 1356–1364.

R Core Team. (2017). *R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2017*. ISBN3-900051-07-0 https://www. R-project. org.

Riley, G. D. (2009). *SSI-4 Stuttering Severity Instrument*. Pro-Ed, Inc.

Sheehan, J. G. (1974). Stuttering behavior: A phonetic analysis. *Journal of Communication Disorders*, *7*(3), 193–212.

Stockwell, R. G. (2007). Why use the S-transform. *Pseudo-Differential Operators: Partial Differential Equations and Time-Frequency Analysis*, *52*, 279–309.

Taylor, I. K. (1966). The properties of stuttered words. *Journal of Verbal Learning and Verbal Behavior*, *5*(2), 112–118.

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

Wessel, J. R., & Aron, A. R. (2017). On the globality of motor suppression: Unexpected events and their influence on behavior and cognition. *Neuron*, *93*(2), 259–280.

Yada, Y., Tomisato, S., & Hashimoto, R. (2018). Online cathodal transcranial direct current stimulation to the right homologue of Broca's area improves speech fluency in people who stutter. *Psychiatry and Clinical Neurosciences*.

Figure Legends

Fig 1. Behavioral task. Each trial began with a fixation cross of variable duration (Baseline period). Stimulus words appeared in the center of the screen followed by a blank screen of variable duration. For speak trials, a white asterisk appeared (cue), signaling the requirement to speak the word on the following green asterisk (go signal). Participants had 2 – 3 s to produce the words. Catch trials started in the same manner, however, a red asterisk appeared after the initial blank screen, indicating that participants should remain silent until the next trial.

Fig 2. Time frequency analysis. A. Greater beta power prior to stuttered compared to fluent speech. The plot shows the difference (*stuttered* minus *fluent*) time-frequency plot with the significant cluster in the beta band highlighted (black contour). Significance was determined via a cluster-based permutation test ($p < 0.05$; two-tailed; 1,024 permutations) across participants (N = 29). The time-frequency analysis window spanned 1 second between the cue (white asterisk; Fig. 1) and the go signal (green asterisk). Zero (0) is the time when the cue was presented while 1 is the time when the go-signal was presented. **B. Correlation between beta power and stuttering severity.** Scatterplot of the relationship between SSI-4 scores (a proxy for stuttering severity) and beta power (22 to 27 Hz, 0.2 to 0.5 seconds after the cue). Spearman rho values and corresponding *p* values reported in the plot for the correlation without (green) and with (red) the outlier (red dot at lower right).

Fig 3. Greater beta power in the R-preSMA prior to stuttered speech. Significant cluster shown in a lateral (**left**), medial (**middle**), and axial (**right**) views. Significance was determined via a cluster-based permutation test across participants (one-tail) using each participant's power spectral density estimates in source space (5124 sources) averaged over frequencies 22 – 27 Hz and times 0.1 to 0.5 sec after the cue, with a *t*-threshold of $p < 0.01$ and a subsequent cluster threshold of $p < 0.05$. Timeline shows pertinent window of interest in orange.

Fig 4. R-preSMA beta power predicts the percentage of trials stuttered. Left: Linear regression between beta power changes from baseline in the R-preSMA and percentage of trials stuttered. The plot shows a positive linear relationship both with (solid black line) and without

(dotted red line) the 2 outliers (red dots; *Materials and methods*). **Right:** Exponential fit to the same data as the left plot, with (solid black line) and without (dotted red line) the outliers. The greater variance explained (R^2 values) for the exponential fit suggests that the percentage of trials stuttered grows exponentially with beta power increases from baseline.

Fig 5. Whole-brain ERF analysis results. The analysis yielded four distinct clusters with greater activity in stuttered compared to fluent trials ($p < 0.05$). Between 200 and 300 ms after the Cue, there was greater activity for stuttered trials in the right dorsal frontal cortex (red cluster), in a cluster largely overlapping in time and spatial location that identified in the PSD analysis (inset: preSMA cluster from Fig. 3 [magenta]; overlap [yellow]). This was followed by greater activity for stuttered trials between 500 and 600 ms in the left superior frontal gyrus (green cluster) and in the right dorsolateral prefrontal cortex (blue cluster). There was also greater activity for stuttered trials in the left posterior superior and middle temporal cortex (dark violet cluster) between 550 and 650 ms. The timeline of events in the experimental trial and the temporal segment corresponding to the epoched data (0 to 1 second after Cue presentation) are shown below the clusters. The approximate times in which significant differences were found are indicated by horizontal bars with the color of each cluster. RH: Right Hemisphere. LH: Left Hemisphere.

Study	N	Task and % of utterances stuttered	Technique	Results	
				stuttered>fluent	**stuttered<fluent**
Braun et al. (1997)	18 (8F)	Narrative speech*; sentence construction*	PET	anterior forebrain (prefrontal/preCG)	post-Rolandic sensory regions
Fox et al. (2000)	10 M	Reading paragraphs (62%)[$]	PET	preCG, SMA, L-IFG	STG; MTG
Ingham et al. (2004)	10 F	Reading paragraphs (73%)[$]	PET	L-occipital	R-STG, L-IFG,R-prefrontal, R-parietal
Sowman et al. (2012)	1 F	Vowel production (58%)	MEG	R-orbitofrontal and inferior-frontal cortices	L-orbitofrontal and inferiorfrontal cortices
Jiang et al. (2012)	20 M	Sentence completion*#	fMRI	L-IFG, bilateral precuneus	
Wymbs et al. (2013)	4 M	Reading words (50%)	fMRI	L-preCG, R-SMA, L-precuneus	
den Ouden et al. (2013)	1 M	Reading words (54%)	fMRI	R-preCG, L-MFG, L-SFG	R-IFG, L-MFG, R-precuneus, L-occipital, L-temporal
Vanhoutte et al. (2016)	7 (3F)	Picture naming*	EEG		
Mersov et al. (2018)	8[@]	Reading words (17%)	MEG		
Connally et al. (2018)	8 (1F)	Sentence reading; picture description*#	fMRI	bilateral IFG, premotor cortex	no statistically significant differences
Korzeczek et al. (2022)	13	Produced pseudowords in a cued-reading paradigm	EEG		

Table 1. Studies that include a state-level comparison (i.e., stuttered versus fluent). Percentages indicate % stuttered utterances across entire sample. Empty cells indicate no statistically significant differences. *percentage of utterances stuttered not reported. [$]Window-interval method used by which 4 s windows were determined to be stuttered, despite including fluent speech. [#]non-stuttering-like disfluencies included as stuttered. [@]gender not reported.

ID	% Trials stuttered	% Trials stuttered (independently-generated)	# Trials	% Trials stuttered (researcher-assisted)	# Trials	SSI-4	SSI-4 severity	Q1	Q2	Q3	Reaction time (ms)
1607	46%	47	170	43	81	22	Mild	1.43	3.57	4.29	258.62
1609	63%	100	3	62	326	30	moderate	3.57	3.57	2.86	289.65
1610	67%	100	5	57	279	20	mild	0.71	0.71	1.43	383.42
1611	78%	89	9	77	243	32	severe	4.29	3.57	2.14	
1613	22%										473.85
1614	32%					19	mild	2.14	2.86	3.57	240.26
1615	48%					31	moderate	2.14	2.86	4.29	737.81
1619	10%					16	very mild	2.14	3.57	5.00	
1620	5%					26	moderate	2.86	3.57	2.14	
1621	68%					30	moderate	2.86	2.14	1.43	
1634	14%	0	1	15	288	37	very severe				700.10
1636	65%	100	2	65	290	28	moderate	5	4	4	344.04
1637	7%	30	23	5	272	19	mild	2	2	3	643.71
1675	50%					35	severe	4	4	2	711.11
1678	3%	0	15	3	237	10	very mild	2	3	2	448.16
1679	5%	9	11	5	238	10	very mild	1	3	3	634.70
1681	6%	13	23	12	258	13	very mild	1	1	2	390.76
1686	48%	50	4	48	204	17	very mild	1	3	5	1071.84
1687	36%	63	35	32	234	18	mild	3	4	5	833.58
1690	33%	11	9	34	278	25	moderate	2	1	1	658.06
1695	44%	100	5	43	284	9	very mild	3	3	2	764.17
1700	92%			92	246	35	severe	3	4	4	1264.72
1705	44%			44	288	18	mild	2	3	4	692.76
1709	48%	82	17	46	276	11	very mild	3	3	2	686.09
1722	62%	48	42	65	250	26	moderate	3	3	3	669.07
1733	50%	61	18	49	273	19	mild	2	2	4	808.89
1734	25%	42	12	24	276	34	severe	2	3	4	532.59
1736	31%	83	6	29	259	13	very mild	2	3	3	1040.48
1737	6%	0	17	7	250	15	very mild	3	4	3	893.65
1738	32%	82	11	30	279	24	mild	3	3	3	1265.39
1742	32%	93	44	18	228	19	mild	1	3	4	810.22

Table 2. Behavioral data per participant. Percent (%) trials stuttered includes all trials except catch trials and errors. Independently-generated words include anticipated words identified by participants without any prompting from the researcher; researcher assisted anticipated words included those identified with researcher prompts. These words were coded a posteriori by reviewing the videos for visit. These videos were not available for participants 1613, 1614, 1615, 1619, 1620, or 1621. Q1: How severe would you rate your stuttering?; Q2: How severe would other people rate your stuttering?; Q3: Overall, how much does stuttering impact your life? (1 = Mild, 5 = Severe). ms = milliseconds. Due to researcher error, participants 1607-1621 completed a 7-point Likert scale for Q1-Q3, which was subsequently normalized to a 5-point scale.