Journal of Speech, Language, and Hearing Research

Conflicting evidence for a motor timing theory of stuttering: choral speech changes the rhythm of both neurotypical and stuttering talkers, but in opposite directions
--Manuscript Draft--

Manuscript Number:	JSLHR-24-00405R1
Full Title:	Conflicting evidence for a motor timing theory of stuttering: choral speech changes the rhythm of both neurotypical and stuttering talkers, but in opposite directions
Article Type:	Research Article
Section/Category:	Speech—CMS2024
Funding Information:	Royal Society (DHF\R1\211078) — Dr Sophie Meekings Economic and Social Research Council — Dr Sophie Meekings
Keywords:	stuttering; motor timing; speech rhythm
Manuscript Classifications:	Fluency Disorders; Stuttering; Speech Production

Abstract:

Purpose

Talking in unison with a partner, otherwise known as choral speech, reliably induces fluency in people who stutter. This effect may arise because choral speech addresses a hypothesised motor timing deficit by giving people who stutter an external rhythm to align with and scaffold their utterances onto. This study tested this theory by comparing the choral speech rhythm of people with and without a stutter to assess whether both groups change their rhythm in similar ways when talking chorally.

Method

20 adults with a stutter and 20 neurotypical controls read a passage on their own, and then a second passage chorally with a neurotypical partner. Their speech rhythm was evaluated using Envelope Modulation Spectrum analysis to derive peak frequency, a measure of the dominant rate of modulation in the sound envelope across several octave bands associated with different parts of speech.

Results

The two groups displayed opposing patterns of rhythmic change during choral reading. People with a stutter increased their EMS peak frequency when they read chorally, while neurotypical talkers' choral speech was characterised by reduced peak frequency compared to solo reading.

Conclusions

Our findings show that the choral speech rhythm of people who stutter differs from that of neurotypical talkers. This indicates limited support for the hypothesis that choral speech addresses a motor timing deficit by giving people who stutter a rhythmic cue to align with.

Response to Reviewers:	Many thanks to both reviewers for their useful and insightful comments, as well as their kind words on the scientific merit of the paper. In general, we agree with all of the points raised and have incorporated their suggestions into the revised manuscript. A more detailed point-by-point response follows. Response to reviewers: Reviewer #1: Reviewer 1: This is an important paper on how choral speaking increases fluency in stuttering and whether this fluency-enhancing strategy matches with hypothesized motor timing mechanisms. I have some comments to improve the clarity of the paper. • Are there any more recent citations on other fluency enhancing strategies other than

the citations from the 1950's?
RESPONSE: We have added some additional citations.
• I recommend using "people who stutter" and then shortening it to PWS for the rest of the paper. For example, "stuttering talkers" in line 73 is an inappropriate term to use (person-first language is required for this journal) so replacing that with PWS would work well.
RESPONSE: We have made this change throughout the paper.

• Lines 76 through 78, I agree with the first two points, but why should speaking chorally change the rhythm to match that of control speakers? This needs to better justified. Line 84 provides a justification but it's not strong enough.
RESPONSE: We have added the following additional text:
When neurotypical talkers speak chorally, their speech also changes in relatively consistent ways that make their rhythm more predictable, for example by regularising vocalic interval durations (Cerda-Oñate et al., 2021; Cummins, 2009). If people who stutter use their partner's speech as a rhythmic guide during choral reading, then it is to be expected that their speech rhythm will also become more predictable, similar to that of typical talkers during choral speech.

• Line 89- EMS in choral speaking did not predict fluency during choral speaking or fluency in spontaneous speech? Also, this claim seems to be in contrast to the claim in line 115.
RESPONSE: We have clarified as follows:
Additionally, Meekings et al. (2023) found that while one measure of rhythm (EMS peak frequency) correlated with induced fluency, the 'rhythmic signature of choral speech' – a combination of EMS metrics that most reliably characterised the choral speech of people who stutter– did not significantly predict fluency in choral speech.
• The stats and acoustic analyses are well done and robust
RESPONSE: Thank you!
• Figure captions need to be more detailed to explain what the boxplots mean (dark line, quartile boxes, dots, colors, etc)
RESPONSE: We have done this.
• Can you provide an r squared value between stuttering frequency and EMS peak frequency?
RESPONSE: Yes- now included.

• Line 291 is a really interesting interpretation and makes sense with the results. A follow-up study could look at a choral reading with a recording to read with to see how EMS values change between neurotypical and stuttering group. I think that is what line 293 is saying but it is confusing.
RESPONSE: That is indeed what we meant- have rephrased the line slightly to make this clearer.
Reviewer #2:
Reviewer 2: The study aims to test whether choral speech, a known fluency enhancer for people who stutter, works by giving the speaker an external rhythmic cue to time their speech to. To do this, the study records people who stutter and neurotypical controls as they read a passage with, and without, a choral speech partner. The outcome measure was the peak frequency using an Envelope Modulation Spectrum analysis. Overall, people who stutter had lower peak frequencies (~= faster speech) during solo speech compared with choral speech, whereas controls had higher frequency speech during solo speech compared with the controls, which was then slower during choral speech. The authors suggest that this is evidence that people who stutter do not adjust their speech rhythm to model that of the neurotypical partner during choral speech. This would suggest that people who stutter are not using their partner's rhythm as a way to increase fluency and that another, unknown, mechanism must explain the fluency enhancing effect.

The study is very clearly written with all necessary details to enable replication attempts. The study is well motivated and builds from prior literature. My main concern is whether the research question can be adequately answered based on the design of the study. The authors do note a few of the following concerns in the discussion.

1) The study specifically aims to test a motor timing theory of stuttering. There is a central assumption that people who stutter will move towards the rhythm of the

neurotypical speaker because the person who stutters has atypical rhythm and therefore wants to take cues from the neurotypical speaker (lines 83-85). However, during choral speech, both pairs are likely to adapt their speech, which is in accordance with the motor timing hypothesis (i.e. both partners settle on an agreed rhythm early on and then it becomes consistent and predictable for the person who stutters) and the results of this study (neurotypical speakers adapt their speech during choral reading). However, in the stuttering-neurotypical pairing, the results only show the frequency of the person who stutters. Therefore, it cannot be ruled out that the neurotypical partner was speaking at a higher peak frequency and therefore people who stutter do in fact match the neurotypical speaker. It could be interesting to plot the peak frequency of the individual pairs of neurotypical speakers, to see if they do converge to each other and in which direction, though this still will not address the central question.

RESPONSE: We agree that this is a limitation of the study, as discussed in l.285-294. We have edited that section and moved it to the top of the limitations section to make this clearer, as follows:

Because only one side of the interaction was recorded in the original experiment with PWS, we were unable to compare their speech rhythm directly with that of their partner, meaning that we cannot assess the degree or directionality of convergence in the PWS-experimenter dyads. Instead, we have compared the choral speech behaviour of typical talkers and PWS when paired with a neurotypical partner. Looking at participants' behaviour independent of that of their partner allowed us to control for the possibility that talkers' speech converges during choral reading not because the person with a stutter is converging with the neurotypical talker, but because the neurotypical talker is converging with the person with a stutter. Previous research has found that neurotypical speakers make rhythmic changes during choral speech which make the perceptual centres of their speech more predictable, and that these changes are typically consistent between participants (Cummins, 2009), so we expect neurotypical participants and experimenters to use approximately the same speech rhythm when reading the same passage chorally. However, it may be that typical speakers adopt a different rhythmic strategy when reading chorally with someone who stutters, or that the specific experimenters who partnered with PWS in this study had higher than average peak frequency, meaning that the PWS did in fact entrain to their partner.

We have also plotted neurotypical within-pair convergence and included this in the Results section- thank you for this suggestion!.

2) The issue, in my opinion, is further compounded by differences in the set-up of the choral partners between people who stutter and controls: people who stutter were paired with an experimenter, who likely has knowledge of stuttering and choral reading, and who has likely read the passage many times before and may have spoken the participant before the start of the experiment(?). The control group, on the other hand, were paired with another neurotypical control and were both reading the passage for first time, presumably without any knowledge of choral reading or stuttering. These factors could have quite a significant effect on the results from the perspective of social convergence theory (e.g. Gregory & Webster, 1996 https://psycnet.apa.org/fulltext/1996-01769-009.html) or simple practice effects – see recent paper from Bradshaw & McGettigan for a nice summary (https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0258747)

RESPONSE: This is an interesting and valid point. Research into synchronous speech suggests that experience with choral reading tends not to affect participants' behaviour, and even practising with the same passage only has a marginal effect. However, it is certainly a possibility that the experimenters' experience with the task combined with their higher social status may have driven them to align less with their partner. We have added a paragraph in the discussion relating to this, as follows:

One experimental difference which may have affected the results is the choice of partner. Neurotypical participants were paired with other neurotypical participants, who were of equivalent social status and had limited previous experience of choral speech. Participants with a stutter were paired with one of four experimenters, who had read the passage before and had more experience with choral reading. Previous research has found that, typically, participants are able to read chorally in close synchrony almost immediately, even if they have no prior experience of the task. Moreover, practising a passage only marginally changes choral speech behaviour, while prior experience with choral speech does not affect performance (Cummins, 2003; O'Dell et

al., 2010). However, because the experimenter might have been perceived as the dominant partner because they ran the experiment and had experience doing the task, this may have caused the PWS to converge more to the experimenter than vice versa (Gregory & Webster, 1996).

3) The results, including statistics and figures, are presented clearly. Consideration could be made to a different interpretation: People who stutter have a lower peak frequency during solo reading because they are employing learned or automatic techniques to control stuttering, such as soft onsets, breath control, elongation of initial sounds, or simply slowed speech pace. When they are then in a fluency enhanced condition, they are able to speak with a higher frequency because they no longer need to use techniques. This interpretation would be held whether or not they are using rhythm as a guide or if something else enables them to gain fluency. It also critically depends on the frequency of the experimenter (see point above). The controls on the other hand, do not need to use techniques when speaking in the solo condition, meaning they can speak at a higher frequency compared with people who stutter. When they speak in the choral condition, they slow down to aid the speech matching process. Within-subject plots and/or within-pair plots may help to untangle this.

RESPONSE: This is an interesting potential explanation, which we have added to the discussion- thank you!

4) Finally, I understand the rationale for wanting to pick one clear measurement which is based on previous research. However, peak frequency alone, in my opinion, cannot answer the central research question without knowing whether each person in the pair converges towards each other and how consistent the frequency then becomes. If people who stutter match with the choral partner and consistently hit the same rhythm, this would answer the main question. A consistency measure could help to answer whether people who stutter maintain a predictable rhythm (even if this information is missing for the choral partner).

RESPONSE: To look at rhythmic consistency, we have included an analysis of peak amplitude, which represents the dominance of the peak frequency. As you'll see, this wasn't super conclusive… but we agree that consistency of speech rhythm could well be an important factor here, and have added a note in the discussion that this should be an avenue of investigation for future studies.

Minor: Can the reasons for the technical difficulties be listed, particularly if this will help people avoid issues in the future?
RESPONSE: This is explained in the Preprocessing section (l. 196-199)- I have additionally amended l. 146 to specify 'issues with recording quality' rather than 'technical issues'.

Minor: Can more details be added from Meekings (2023) justifying why peak frequency is a good measure, what the other measures were and why they were not suitable/chosen for the current project. It is only mentioned that they didn't correlate with stuttering frequency (lines 88, 116).
RESPONSE: We have added the following passage:
Meekings et al. (2023) also identified six EMS metrics that, in combination, significantly predicted whether a participant was speaking chorally or not. These metrics were peak frequency in the 125Hz, 1kHz and 2kHz octave bands, energy between 4-10Hz in the 4kHz band, energy between 3-6Hz in the full band, and ratio of the energy above and below 4Hz in the 8kHz band. Although the weighted linear combination of these variables that predicted choral speech rhythm did not significantly predict fluency, these results do suggest that EMS peak frequency across multiple bands is an important component of choral speech rhythm.

Corresponding Author:	Sophie Meekings University of York York, North Yorkshire UNITED KINGDOM
Other Authors:	Lotte Eijk

	Stefany Stankova
	Santosh Maruthy
	Sophie Kerttu Scott

Click here to access/download;Manuscript;REVISED- Motor Timing JSLHR.docx

Conflicting evidence for a motor timing theory of stuttering: choral speech changes the rhythm of both neurotypical and stuttering talkers, but in opposite directions

Sophie Meekings[1*], Lotte Eijk[1], Stefany Stankova[1], Santosh Maruthy[2], Sophie Kerttu Scott[3]

[1] Department of Psychology, University of York, York, UK

[2] Department of Speech-Language Sciences, All-India Institute of Speech and Hearing, Mysore, Karnataka, India

[3] Institute of Cognitive Neuroscience, University College London, London, UK

[*] Corresponding author: sophie.meekings@york.ac.uk

Conflicts of interest

The authors have no conflicts of interest to declare

Abstract

Purpose: Talking in unison with a partner, otherwise known as choral speech, reliably induces fluency in people who stutter. This effect may arise because choral speech addresses a hypothesised motor timing deficit by giving people who stutter an external rhythm to align with and scaffold their utterances onto. This study tested this theory by comparing the choral speech rhythm of people who do and do not stutter to assess whether both groups change their rhythm in similar ways when talking chorally.

Method: 20 adults who stutter and 20 neurotypical controls read a passage on their own, and then a second passage chorally with a neurotypical partner. Their speech rhythm was evaluated using Envelope Modulation Spectrum analysis to derive peak frequency, a measure of the dominant rate of modulation in the sound envelope, as well as peak amplitude (the amplitude of the peak frequency), across several octave bands associated with different features of speech.

Results: The two groups displayed opposing patterns of rhythmic change during choral reading. People with a stutter increased their EMS peak frequency when they read chorally, while neurotypical talkers' choral speech was characterised by reduced peak frequency compared to solo reading.

Conclusions: Our findings show that the choral speech rhythm of people who stutter differs from that of neurotypical talkers. This indicates limited support for the hypothesis that choral speech addresses a motor timing deficit by giving people who stutter a rhythmic cue to align with.

Keywords

stuttering; motor timing; speech rhythm

Introduction

Persistent developmental stuttering is a lifelong neurological condition affecting around 1% of adults worldwide (Yairi & Ambrose, 2013). This condition, which is characterised by syllable repetitions, prolongations, and silent 'blocks' during speech, can cause communication difficulties with potentially far-reaching effects on wellbeing and quality of life (Craig et al., 2009). Although there is no 'cure' for stuttering, people who stutter (PWS) do experience periods of relative fluency (Bloodstein, 1949; Bloodstein et al., 2021). While these periods are often spontaneous and unpredictable, many PWS report that they are also reliably more fluent in specific situations (Bloodstein, 1950; Budde et al., 2014). Understanding which situations induce fluency, and the mechanism by which they do so, may provide insights into the aetiology of stuttering and inform therapies.

Fluency-enhancing situations vary considerably from person to person, with participants reporting reduced stuttering in conditions ranging from 'speaking to an animal' to 'when under fire during the war' (Bloodstein, 1950). However, researchers have noted that many tactics that are commonly observed to reduce stuttering induce a different speech rhythm. For example, a large body of evidence demonstrates reliable increases in fluency when people who stutter speak in time with a metronome (Hutchinson & Norris, 1977; Toyomura et al., 2011), another speaker (Rami & Diederich, 2005; Saltuklaroglu et al., 2009), or other rhythmic activities such as walking, arm swinging or foot tapping (Andrews et al., 1982; Barber, 1939; Bloodstein, 1950).

This evidence has led some researchers to suggest that stuttering is a manifestation of an underlying difficulty with the initiation and timing of complex rhythmic motor gestures (Etchell et al., 2014; Max et al., 2003). Rhythmic stimuli thus act as a 'pace-setting' mechanism which induces fluency by introducing an external signal to align with, so that the talker need not struggle to generate their own rhythmic patterns (Dechamma & Maruthy, 2018; Pattie & Knight, 1944). This theory appears to be supported by research demonstrating that PWS show increased variability in rhythmic movement such as finger tapping (Sares et al., 2019; Slis et al., 2023) and differences in lip, jaw and vocal tract movement (McClean et al., 2004; Wiltshire et al., 2021) compared to typical speakers, although other work has found no difference between PWS and controls for simple isochronous rhythm production tasks (Max & Yudman, 2003).

Choral speech, in which talkers read or recite words in unison, is known to be among the most effective ways to induce this 'rhythm effect' on fluency (e.g. Andrews et al., 1982; Barber, 1939; Kalinowski & Saltuklaroglu, 2003). Reading chorally immediately and dramatically reduces stuttering frequency and duration by between 90-100% compared to solo speech

(Dechamma & Maruthy, 2018; Kiefte & Armson, 2008; Meekings et al., 2023), and is more effective than other commonly used fluency enhancing manipulations such as altered auditory feedback (Kiefte & Armson, 2008) or speaking in time with other rhythmic stimuli (Barber, 1939a; Bloodstein, 1950). This therefore offers an ideal paradigm with which to investigate the 'pace-setting' theory.

If PWS experience difficulties generating and sequencing motor speech gestures, and speaking chorally induces fluency because it provides an external signal for PWS to synchronise with, then:

1. The speech rhythm of PWS should change when they read chorally
2. This change in rhythm should be correlated with increased fluency
3. The change in rhythm should also match that of typical controls

Previous work supports the first two elements of this hypothesis: choral reading changes the speech rhythm of PWS, and this change in rhythm is associated with improvements in fluency (Meekings et al., 2023; Dechamma & Maruthy, 2019). When neurotypical talkers speak chorally, their speech also changes in relatively consistent ways that make their rhythm more predictable, for example by regularising vocalic interval durations (Cerda-Oñate et al., 2021; Cummins, 2009). If people who stutter use their partner's speech as a rhythmic guide during choral reading, then it is to be expected that their speech rhythm will also become more predictable, similar to that of typical talkers during choral speech.

However, studies into the choral speech rhythm of PWS have not evaluated the choral speech rhythm of a neurotypical control group, while studies into neurotypical choral speech have typically used duration-based measures that may be skewed by disfluencies, meaning that it is difficult to compare these results to the population of PWS. Additionally, Meekings et al. (2023) found that while one measure of rhythm (envelope modulation spectrum peak frequency) correlated with induced fluency, the 'rhythmic signature of choral speech' – a combination of acoustic metrics that most reliably characterised the choral speech of PWS– did not significantly predict fluency in choral speech. It thus remains unclear whether these changes in speech rhythm are the result of copying neurotypical talkers, or whether choral speech induces fluency by some other means, which then causes differences in speech rhythm (since fluent speech is by definition rhythmically different from stuttered speech). To establish a causal link, it is necessary to compare the choral speech behaviour of PWS with that of neurotypical controls.

In this paper we use Envelope Modulation Spectrum (EMS) analysis to characterise speech rhythm during choral reading in PWS and neurotypical controls. The envelope of speech contains multiple regular amplitude fluctuations corresponding to, for example, patterns of word stress, or syllable durations. EMS analysis quantifies the periodicity of these slow modulations in the amplitude envelope of speech by Fourier transforming the envelope to derive a "rhythm spectrum" that has peaks in frequencies that correspond to repeating patterns (Tilsen & Johnson, 2008). These spectra are computed for the full signal and for octave bands centred around frequencies from 125 to 8000 Hz; each band contains independent information about rhythmic patterns at different levels of the speech signal, such as syllables, vowels or bursts and fricatives (Crouzet & Ainsworth, 2001; Liss et al., 2010), although the exact correspondence between envelope measures and phonetic features is still under investigation (MacIntyre et al., 2022).

In contrast to duration-based measures of rhythm, such as speech rate, EMS metrics are not skewed by characteristic features of atypical speech such as long silent pauses (White & Mattys, 2007), meaning that the analysis can be applied to the whole speech signal with no need to remove potentially significant disfluencies (Liss et al., 2010). Previous research has used this approach to investigate atypical speech patterns in people with dysarthria (Liss et al., 2010), apraxia of speech (Basilakos et al., 2017) and PWS (Dechamma & Maruthy, 2018; Meekings et al., 2023).

There are several different metrics that can be derived from the EMS, but here we focus on peak frequency and peak amplitude. The 'rhythm spectrum' generated by EMS analysis contains multiple peaks, each representing the frequency at which a particular amplitude modulation repeats itself. The peak with the highest amplitude is the most periodic pattern. Measuring the frequency at which that peak occurs tells us the overall dominant rate of modulation in the signal, while its amplitude tells us how dominant that pattern is. These metrics have been used in previous research on this subject (Dechamma & Maruthy, 2019; Meekings et al., 2023). Specifically, these studies found that choral speech in PWS was associated with an increase in both fluency and mean peak frequency across several octave bands compared to solo speech. Dechamma & Maruthy additionally found that PWS' mean peak amplitude was lower during choral speech than solo reading. Meekings et al. (2023) also identified six EMS metrics that, in combination, significantly predicted whether a participant was speaking chorally or not. These metrics were peak frequency in the 125Hz, 1kHz and 2kHz octave bands, energy between 4-10 Hz in the 4kHz band, energy between 3-6Hz in the full band, and ratio of the energy above and below 4Hz in the 8kHz band. Although the weighted linear combination of these variables that predicted choral speech rhythm did not

significantly predict fluency, these results do suggest that EMS peak frequency across multiple bands is an important component of choral speech rhythm.

To sum up, this study aims to evaluate the theory that PWS become more fluent when they speak chorally because they adopt the same choral speech patterns as neurotypical talkers. If this is true, we expect to see both groups' speech rhythm (measured using EMS mean peak frequency) become more similar during choral reading compared to solo speech. If, however, rhythmic changes are caused by fluency rather than vice versa, we might expect the choral speech rhythm of PWS to look more similar to the fluent speech of neurotypical talkers when they speak non-chorally.

Even if PWS and neurotypical controls use different speech rhythms when talking chorally, it may be the case that the choral speech task induces more consistent speech rhythm in both partners, which in itself is an aid to fluency. To assess this, we look at changes in peak amplitude: we expect speech that is more consistent in its rhythm to have a higher peak amplitude.

Method

Ethical approval for this study was granted by the UCL Psychology Ethics Committee for the group of PWS (approval ID: ICN-PWB-13-12-13) and the University of York Psychology Ethics Committee for the neurotypical controls (approval ID: 2212).

Participants

PWS and neurotypical participants were recruited and tested separately at University College London and the University of York respectively. All participants were adult British English speakers with normal hearing.

Twenty-five PWS (17 male, 8 female; mean age 37.5) were recruited through STAMMA, the British Stammering Association, as part of a larger neuroimaging study (Meekings et al., 2020). They were paired with one of four experimenters to complete the choral speech task; the experimenters' voice was not recorded. All participants were adults who self-identified as a PWS. Their speech was assessed using Riley's Stuttering Severity Instrument IV (SSI-IV; Riley, 1975) and participants were included in the final analysis if they had a stutter of any severity as defined by the SSI-IV. Their SSI-IV scores ranged from 6 to 44 out of a possible 46 (M=23.4).

Thirty-six neurotypical controls (27 female; 7 male; 2 non-binary, mean age 20.3) were recruited through the University of York and completed the task in pairs; both partners' voices were recorded.

Participants and choral speech pairs were not controlled for gender as this has not been found to influence choral speech behaviour (Poore & Ferguson, 2008). Hearing threshold was assessed using an Amplivox 116 Screening Audiometer with Audiocup headphones. Participants were included if their binaural four-frequency pure tone average hearing threshold was 20dB HL or less. One participant with a stutter did not meet this threshold and was excluded from the analysis. All other participants had normal hearing. An additional four PWS and eight neurotypical pairs' data was excluded due to recording quality issues. More details on the reasons for exclusion are given in the Analysis section below. Thus, 20 neurotypical participants (15 female, 3 male, 2 non-binary) aged between 18-26 (M = 20.1, SD = 1.8) and 20 adults who stutter (8 female, 12 male) aged between 18-61 (M = 38.4, SD = 12) were included in the final analysis.

Procedure

Participants provided informed consent and completed a hearing screening as described above. They were then seated in a sound-attenuated booth to complete the reading-aloud tasks. At the end of the experiment they were debriefed and asked to complete a short demographics questionnaire to collect information about age and gender. The experiment took around 40 minutes in total.

Stimuli for the experiment were taken from Riley's Stuttering Severity Instrument (Fourth Ed.) and consisted of two passages from the adult reading materials, adapted from a travel journalism article (Riley, 1972). Passage A was 374 syllables and passage B was 369 syllables. Participants were asked to read one of the passages on their own, and the second in unison with their partner. The order of the conditions was always the same (solo followed by choral reading) to avoid any spillover effects on speech rhythm from the choral condition. However, the passage presented in each condition was pseudorandomised such that half the participants read Passage A in the solo condition and Passage B in the choral condition, and half vice versa.

Before beginning each task, participants were instructed to read the given passage through, ask the experimenter if they had any questions about word pronunciation, and then signal when they were ready to begin. The experimenter then began the recording and gave a visual or verbal signal for the participant to begin reading aloud. Participants were told to 'keep

together as much as possible' during the choral condition but were otherwise not given specific instructions on how to complete the tasks.

Technical specifications

During the choral condition, all participants were visually separated from their partner and heard each other through headphones. PWS and neurotypical participants were tested separately at different institutions with conditions kept as similar as possible across the groups within the constraints of the available space and equipment.

PWS sat inside a sound-attenuated booth and spoke into a RODE NT1-A one-inch cardioid condenser microphone connected to a Windows computer via a Fireface UC audio interface (RME Audio, Haimhausen). They heard their partner through Beyerdynamic DT 770 Pro closed-back circumaural headphones. Their voices were recorded in stereo at 44.1kHz with 16 bit quantisation using Adobe Audacity 3.0. Their choral speech partner was positioned outside the testing booth, spoke into an AKG 190E cardioid dynamic microphone connected directly to the booth sound system (i.e. not routed through a computer and thus not recorded) and heard their partner through AKG K240 Studio on-ear headphones.

Neurotypical pairs were seated in a sound-attenuated booth with a dividing screen between them, wearing Beyerdynamic DT 770 Pro headphones. They spoke into Sennheiser HSP-4 cardioid condenser microphones connected to a Windows computer via an RME Fireface UFX II audio interface. Audio was recorded in stereo (with one participant in each channel) at 44.1kHz with 16 bit quantisation using Audacity 3.4.2.

Analysis

Preprocessing

Audio files were reviewed in Audacity and were excluded from analysis if either member of a participant pair had recording or sound quality issues that might affect acoustic analyses (for example, audio clipping or frequent nonverbal sounds such as coughing or laughing). Four PWS and eight neurotypical pairs were excluded from analysis for this reason, as detailed in the Methods section.

The remaining audio files were converted from stereo to one mono file per participant with a sampling rate of 32kHz using the tuneR package in R, to meet the file specifications necessary for the EMS analysis. This analysis was conducted on the mono files using a custom script in MATLAB R2023b following the procedure outlined in Liss et al. (2010). The amplitude envelope was extracted from the full-band signal and six octave bands centred around 125, 250, 500, 2000, 4000 and 8000 Hz. These envelopes were half-rectified, low-pass filtered at 30Hz using a fourth-order Butterworth filter, and downsampled to 80 Hz. The power spectra of the resulting envelopes were calculated using the Goertzel algorithm and converted to decibels for frequencies up to 10 Hz. The peak frequency of each power spectrum was measured using the same MATLAB script and entered into the analysis. Analysis was conducted on mean peak frequency values and peak amplitude values collapsed over all octave bands, in keeping with previous research indicating that they are significant predictors of changes between solo and choral reading conditions in the choral speech of people who stutter (Meekings et al., 2023; Dechamma & Maruthy, 2019).

Stuttering utterances in both conditions were additionally evaluated for stuttering frequency, measured as the percentage of syllables stuttered.

Statistical analysis

A linear mixed-effects regression model (1) was constructed to assess differences in EMS peak frequency as a function of condition (choral vs solo speech) and group (neurotypical vs PWS), with a condition by group interaction term. The model included random intercepts for subject and passage to control for rhythmic differences between the passages and individual variation in features such as age and vocal characteristics. Peak amplitude was assessed using a linear model with no random effects (2), as a LMER demonstrated that variance attributable to the two random factors included in model (1) was at or close to zero.

For PWS, fluency (percentage of syllables stuttered while reading each passage) was modelled as a function of EMS peak frequency to evaluate the relationship between rhythmic changes and fluency, and condition to confirm the fluency-inducing effects of choral speech. This model included random intercepts for subject only, as- similarly to the model for peak amplitude- adding random intercepts for passage led to singular model fit owing to a lack of variance.

(1) Peak frequency ~ Condition * Group + (1 | Subject) + (1 | Passage)
(2) Peak amplitude ~ Condition * Group

(3) Fluency ~ Peak frequency + Condition + (1 | Subject)

Analysis was performed using R version 4.3.1 (R Core Team, 2021), with lme4 1.1.34 (Bates et al., 2015) to run the linear mixed effects models and lmerTest 3.1.3 (Kuznetsova et al., 2017) to provide p-values using Satterthwaite's approximation for degrees of freedom. Data processing and visualisation was performed using tidyverse 2.0.0 (Wickham et al., 2019).

Results

EMS peak frequency

When neurotypical dyads spoke chorally, their peak frequencies became more similar. Figure 1 illustrates this pattern, demonstrating that choral speech leads to rhythmic convergence in neurotypical dyads.

[Figure 1]

Since we were unable to assess convergence for the PWS dyads, we instead compared their speech rhythm behaviour with that of the neurotypical dyads in each condition. There was a significant effect of condition (β = -0.17, *SE* = 0.06, *t* = -2.74, *p* = .006) and group (β = -0.27, *SE* = 0.08, *t* = -3.38, *p* < .001), and a significant condition by group interaction (β = 0.39, *SE* = 0.09, *t* = 4.003, *p* < .001). Results are illustrated in Figure 2, shown for each passage separately to allow for the probability that reading different materials evokes different speech rhythms.

[Figure 2]

Consistent with previous research, PWS' EMS peak frequency was higher during choral reading compared to reading alone. Neurotypical participants exhibited the opposite pattern: their peak frequency was higher when they spoke on their own compared to reading chorally. Additionally, during solo reading, PWS' peak frequency was significantly lower than controls': neurotypical controls read the passages with a mean peak frequency of 0.63 overall, compared with 0.36 for participants with a stutter. This pattern reversed when participants spoke chorally: in this case, PWS' peak frequency was significantly higher than controls. During choral speech, neurotypical controls spoke with a mean peak frequency of 0.46, while participants with a stutter read with a mean peak frequency of 0.58.

EMS peak amplitude

Boxplots of the data (Figure 3) suggested that peak amplitude was slightly more variable in the solo condition than in the choral condition for both groups. However, there were no significant effects of condition (β = 0.56, *SE* = 4.38, *t* = 0.13, *p* = .90), group (β = -0.58, *SE* = 5.02, *t* = -0.12, *p* = .91), or any significant condition by group interaction (β = -0.12, *SE* = 7.06, *t* = 0.02, *p* = .99); the adjusted r squared value for the model was -0.004.

[Figure 3]

Stuttering frequency

As previously reported in Meekings et al. (2023), PWS' fluency increased when they spoke chorally (β = 3.96, *SE* = 0.41, *t* = 9.66, *p* < .001). As peak amplitude did not significantly differ between conditions, we did not test for a correlation between this metric and stuttering frequency. However, linear mixed modelling revealed that fluency was significantly predicted by mean peak frequency (β = -1.72, *SE* = 0.42, *t* = -4.07, *p* < .001, conditional r^2 = 0.07, marginal r^2 = 0.02). The relationship between stuttering frequency and mean EMS peak frequency is shown in Figure 4 below.

[Figure 4]

Discussion

Dechamma & Maruthy (2019) and Meekings et al. (2023) both found that choral speech in participants with a stutter was associated with an increase in both fluency and mean peak frequency across multiple octave bands, compared to solo speech. This suggested a relationship between speech rhythm and fluency that might support the hypothesis that PWS use their partner's speech as a rhythmic guide during choral reading. However, our data suggests that the choral speech rhythm of neurotypical talkers is significantly different to that of people who stutter. Moreover, reading chorally affects these populations in opposite directions: peak frequency is higher during choral reading than solo speech in PWS, but the reverse is true for neurotypical talkers. The choral speech rhythm of PWS is similar to the solo speech of neurotypical controls, but very different from neurotypical choral speech.

We interpret these results as indicating that fluency induced during choral reading may not be caused by imitating typical speakers' choral vocal patterns. That is, the change in speech rhythm does not cause the change in fluency; rather, the change in fluency causes the change

in speech rhythm. This could be, for example, because PWS use learned or automatic fluency techniques when talking on their own, and these techniques (for example, sound elongation, soft onsets, or slowed speech rhythm) result in a lower EMS peak frequency. When fluency is induced during choral speech, there is no longer any need to use these techniques and thus EMS peak frequency increases.

However, there are some limitations to our findings. Because only one side of the interaction was recorded in the original experiment with PWS, we were unable to compare their speech rhythm directly with that of their partner, meaning that we cannot assess the degree or directionality of convergence in the PWS-experimenter dyads. Instead, we have compared the choral speech behaviour of typical talkers and PWS when paired with a neurotypical partner. Looking at participants' behaviour independent of that of their partner allowed us to control for the possibility that talkers' speech converges during choral reading not because the PWS is converging with the neurotypical talker, but because the neurotypical talker is converging with the PWS. Previous research has found that neurotypical speakers make rhythmic changes during choral speech which make the perceptual centres of their speech more predictable, and that these changes are typically consistent between participants (Cummins, 2009), so we expect neurotypical participants and experimenters to use approximately the same speech rhythm when reading the same passage chorally. However, it may be that typical speakers adopt a different rhythmic strategy when reading chorally with someone who stutters, or that the specific experimenters who partnered with PWS in this study had higher than average peak frequency, meaning that the PWS did in fact converge to their partner. Alternatively, the task instructions might prompt participants to speak with greater rhythmicity, even if they do not copy their partner's speech rhythm, and this more consistent rhythm might be sufficient to induce fluency in and of itself. In this experiment, we used EMS mean peak amplitude to assess the degree to which participants used a consistent speech rhythm. We found no significant differences between conditions or groups. This is surprising, as previous research does suggest that choral speech is more predictable rhythmically than solo reading.

Future research could investigate other measures of rhythmic consistency. Additionally, it would be informative to compare choral speech convergence between neurotypical-neurotypical pairs, neurotypical-PWS pairs and contrast both groups' convergence with a live speaker to convergence with recorded speech, to establish to what extent each partner contributes to convergence during choral speech.

A further limitation is that because PWS' data was originally collected as part of a behavioural pre-test for a larger experiment, while the control data was collected later at a different

institution, this led to some small differences in equipment and experimental setup which may have affected the results, although every effort was made to ensure a comparable testing environment and matching technical specifications between sets of equipment.

One experimental difference which may have affected the results is the choice of partner. Neurotypical participants were paired with other neurotypical participants, who were of equivalent social status and had limited previous experience of choral speech. Participants who stutter were paired with one of four experimenters, who had read the passage before and had more experience with choral reading. Previous research has found that, typically, participants are able to read chorally in close synchrony almost immediately, even if they have no prior experience of the task. Moreover, practising a passage only marginally changes choral speech behaviour, while prior experience with choral speech does not affect performance (Cummins, 2003; O'Dell et al., 2010). However, because the experimenter might have been perceived as the dominant partner because they ran the experiment and had experience doing the task, this may have caused the PWS to converge more to the experimenter than vice versa (Gregory & Webster, 1996).

In general, evidence for the motor timing theory of stuttering is mixed. While there are many effective fluency-enhancing situations that appear to work by inducing a regular speech rhythm, many other interventions such as altered auditory feedback and masking noise have no obvious rhythmic effect. Additionally, while many authors have found increased kinematic variability or differences in movement initiation and timing (Frisch et al., 2016; Loucks & De Nil, 2006; Wiltshire et al., 2021) that are reduced when PWS are provided with an external rhythmic stimulus (Franke et al., 2023; Wiltshire et al., 2023), other studies have found no difference between the motor skills of people who stutter and neurotypical controls (Max & Yudman, 2003; Smith & Kleinow, 2000; Zelaznik et al., 1994).

If fluency is not caused by copying neurotypical speech rhythm, it could be that it is instead induced by some other property of choral speech. Another mechanism that has been hypothesised to explain the effect of fluency-enhancing situations including choral speech is that many of them occlude the sound of the talker's voice, preventing them from using auditory feedback to guide their utterances. However, evidence for this theory is also mixed, with studies finding opposing results (Garnett et al., 2022; Meekings et al., 2020).

It is also possible that PWS do converge to their partner's speech rhythm during choral reading, but to a different feature. We chose EMS peak frequency to investigate as this measure is correlated with fluency during choral reading in people who stutter. However,

recent work suggests that neurotypical talkers synchronise vowel onsets closely during choral speech and has identified a specific amplitude envelope correlated with this (MacIntyre et al., 2022), which may provide a better measure to investigate rhythmic behaviour during this task.

Research into conversational alignment more generally demonstrates that while participants do often adapt to their partner's speech behaviour during conversational interaction, there is individual variation in the precise feature that talkers align with (e.g., Eijk, 2023). In our work, there was considerable variance in all participants' speech rhythm during the passage reading in both choral and solo conditions. Stuttering is also a heterogeneous disorder (SheikhBahaei et al., 2023), and this is reflected in our results: participants with a stutter had a wide range of stuttering severities. It is possible that individuals, and individual dyads, adopt different strategies during choral reading. It may therefore be informative to look at individual patterns of speech adaptation between dyads during choral speech.

Overall, however, our results demonstrate that choral reading affects neurotypical talkers and people who stutter in diametrically opposing ways, suggesting that the fluency-inducing effect of choral reading in PWS may not be attributable to participants copying their neurotypical partner's speech rhythms.

Acknowledgments

Author Contributions (CRediT Statement)

Meekings: Conceptualisation, data curation, formal analysis, project administration, investigation, visualisation, funding acquisition, writing- original draft
Eijk: Investigation, project administration, writing- review & editing
Stankova: Investigation, data curation
Maruthy: Conceptualisation, resources, software
Scott: Supervision, investigation

Data Availability Statement

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

References

Andrews, G., Howie, P. M., Dozsa, M., & Guitar, B. E. (1982). Stuttering: Speech pattern characteristics under fluency-inducing conditions. *Journal of Speech and Hearing Research*, *25*(2), 208–216.

Barber, V. (1939a). Studies in the Psychology of Stuttering, XV [1] This study is part of a Ph.D. thesis directed by Dr. Wendell Johnson at the State University of Iowa as part of a program of studies in the psychology of stuttering. *Journal of Speech Disorders*, *4*(4), 371. https://doi.org/10.1044/jshd.0404.371

Barber, V. (1939b). Studies in the Psychology of Stuttering, XV: Chorus-reading as a distraction in stuttering. *Journal of Speech Disorders*, *4*(4), 371. https://doi.org/10.1044/jshd.0404.371

Basilakos, A., Yourganov, G., den, O. D.-B., Fogerty, D., Rorden, C., Feenaughty, L., & Fridriksson, J. (2017). A Multivariate Analytic Approach to the Differential Diagnosis of Apraxia of Speech. *Journal of Speech, Language, and Hearing Research*, *60*(12), 3378–3392. https://doi.org/10.1044/2017_JSLHR-S-16-0443

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

Bloodstein, O. (1949). Conditions under which Stuttering is Reduced or Absent: A Review of Literature. *Journal of Speech and Hearing Disorders*, *14*(4), 295–302.

https://doi.org/10.1044/jshd.1404.295

Bloodstein, O. (1950). A Rating Scale Study Of Conditions Under Which Stuttering Is Reduced

Or Absent. *Journal of Speech and Hearing Disorders*, *15*(1), 29.

https://doi.org/10.1044/jshd.1501.29

Bloodstein, O., Ratner, N. B., & Brundage, S. B. (2021). *A handbook on stuttering* (Vol. 1).

Plural Publishing.

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. In *Brain and*

Language (Vol. 139, pp. 99–107). https://doi.org/10.1016/j.bandl.2014.10.002

Cerda-Oñate, K., Vega, G. T., & Ordin, M. (2021). Speech rhythm convergence in a dyadic

reading task. *Speech Communication*, *131*, 1–12.

https://doi.org/10.1016/j.specom.2021.04.003

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

who stutter. *Journal of Fluency Disorders*, *34*(2), 61–71.

https://doi.org/10.1016/j.jfludis.2009.05.002

Crouzet, O., & Ainsworth, W. A. (2001). On the Various Influences of Envelope Information on the

Perception of Speech in Adverse Conditions: An Analysis of Between-Channel Envelope

Correlation. *Paper Presented at the Workshop on Consistent and Reliable Cues for*

Sound Analysis; Aalborg, Denmark.

Cummins, F. (2003). Practice and performance in speech produced synchronously. *Journal of*

Phonetics, *31*(2), 139–148. https://doi.org/10.1016/S0095-4470(02)00082-7

Cummins, F. (2009). Rhythm as entrainment: The case of synchronous speech. *Journal of*

Phonetics, 37(1), 16–28. https://doi.org/10.1016/j.wocn.2008.08.003

Dechamma, D., & Maruthy, S. (2018). Envelope modulation spectral (EMS) analyses of solo

reading and choral reading conditions suggest changes in speech rhythm in adults who

stutter. *Journal of Fluency Disorders, 58*, 47–60.

https://doi.org/10.1016/j.jfludis.2018.09.002

Eijk, L. (2023). *Linguistic alignment: The syntactic, prosodic, and segmental phonetic levels*

[Doctoral dissertation]. Radboud University Nijmegen.

Etchell, A. C., Johnson, B. W., & Sowman, P. F. (2014). Behavioral and multimodal

neuroimaging evidence for a deficit in brain timing networks in stuttering: A hypothesis

and theory. *Frontiers in Human Neuroscience, 8*.

https://www.frontiersin.org/articles/10.3389/fnhum.2014.00467

Franke, M., Benker, N., Falk, S., & Hoole, P. (2023). Synchronization type matters: Articulatory

timing in different rhythmic conditions in persons who stutter. *Proceedings of the 20th*

International Congress of Phonetic Sciences.

Frisch, S. A., Maxfield, N., & Belmont, A. (2016). Anticipatory coarticulation and stability of

speech in typically fluent speakers and people who stutter. *Clinical Linguistics &*

Phonetics, 30(3–5), 277–291. https://doi.org/10.3109/02699206.2015.1137632

Garnett, E. O., Chow, H. M., Limb, S., Liu, Y., & Chang, S.-E. (2022). Neural activity during solo

and choral reading: A functional magnetic resonance imaging study of overt continuous

speech production in adults who stutter. *Frontiers in Human Neuroscience, 16*.

https://doi.org/10.3389/fnhum.2022.894676

Gregory Jr., S. W., & Webster, S. (1996). A nonverbal signal in voices of interview partners

effectively predicts communication accommodation and social status perceptions. *Journal of Personality and Social Psychology*, *70*(6), 1231–1240. https://doi.org/10.1037/0022-3514.70.6.1231

Hutchinson, J. M., & Norris, G. M. (1977). The differential effect of three auditory stimuli on the frequency of stuttering behaviors. *Journal of Fluency Disorders*, *2*(4), 283–293. https://doi.org/10.1016/0094-730X(77)90032-8

Kalinowski, J., & Saltuklaroglu, T. (2003). Speaking with a mirror: Engagement of mirror neurons via choral speech and its derivatives induces stuttering inhibition. *Medical Hypotheses*, *60*(4), 538–543. https://doi.org/10.1016/S0306-9877(03)00004-5

Kiefte, M., & Armson, J. (2008). Dissecting choral speech: Properties of the accompanist critical to stuttering reduction. *Journal of Communication Disorders*, *41*(1), 33–48. https://doi.org/10.1016/j.jcomdis.2007.03.002

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*, 1–26. https://doi.org/10.18637/jss.v082.i13

Liss, J. M., LeGendre, S., & Lotto, A. J. (2010). Discriminating Dysarthria Type From Envelope Modulation Spectra. *Journal of Speech, Language, and Hearing Research*, *53*(5), 1246–1255. https://doi.org/10.1044/1092-4388(2010/09-0121)

Loucks, T. M. J., & De Nil, L. F. (2006). Oral Kinesthetic Deficit in Adults Who Stutter: A Target-Accuracy Study. *Journal of Motor Behavior*, *38*(3), 238–247. https://doi.org/10.3200/JMBR.38.3.238-247

MacIntyre, A. D., Cai, C. Q., & Scott, S. K. (2022). Pushing the envelope: Evaluating speech

rhythm with different envelope extraction techniques. *The Journal of the Acoustical Society of America*, *151*(3), 2002–2026. https://doi.org/10.1121/10.0009844

Max, L., Caruso, A. J., & Gracco, V. L. (2003). Kinematic analyses of speech, orofacial nonspeech, and finger movements in stuttering and nonstuttering adults. *Journal of Speech, Language, and Hearing Research : JSLHR*, *46*(1), 215–232.

Max, L., & Yudman, E. M. (2003). Accuracy and Variability of Isochronous Rhythmic Timing Across Motor Systems in Stuttering Versus Nonstuttering Individuals. *Journal of Speech, Language, and Hearing Research*, *46*(1), 146–163. https://doi.org/10.1044/1092-4388(2003/012)

McClean, M. D., Tasko, S. M., & Runyan, C. M. (2004). Orofacial movements associated with fluent speech in persons who stutter. *Journal of Speech, Language, and Hearing Research: JSLHR*, *47*(2), 294–303. https://doi.org/10.1044/1092-4388(2004/024)

Meekings, S., Eijk, L., Maruthy, S., & Scott, S. (2023). *Talking chorally alters speech rhythm and induces fluency in people who stutter, but are these things connected?*

Meekings, S., Jasmin, K., Lima, C., & Scott, S. (2020). *Does over-reliance on auditory feedback cause disfluency? An fMRI study of induced fluency in people who stutter* (p. 2020.11.18.378265). bioRxiv. https://doi.org/10.1101/2020.11.18.378265

O'Dell, M. L., Nieminen, T., & Mustanoja, L. (2010). *Assessing rhythmic differences with synchronous speech*. Speech Prosody 2010-Fifth International Conference.

Pattie, F. A., & Knight, B. B. (1944). Why does the speech of stutterers improve in chorus reading? *The Journal of Abnormal and Social Psychology*, *39*(3), 362–367. https://doi.org/10.1037/h0054206

Poore, M. A., & Ferguson, S. H. (2008). Methodological variables in choral reading. *Clinical Linguistics & Phonetics*, *22*(1), 13–24. https://doi.org/10.1080/02699200701601971

Rami, M. K., & Diederich, E. (2005). Effect of reading with reversed speech on frequency of stuttering in adults. *Perceptual and Motor Skills*, *100*(2), 387–393. https://doi.org/10.2466/pms.100.2.387-393

Riley, G. D. (1972). A Stuttering Severity Instrument for Children and Adults. *Journal of Speech and Hearing Disorders*, *37*(3), 314. https://doi.org/10.1044/jshd.3703.314

Saltuklaroglu, T., Kalinowski, J., Robbins, M., Crawcour, S., & Bowers, A. (2009). Comparisons of stuttering frequency during and after speech initiation in unaltered feedback, altered auditory feedback and choral speech conditions. *International Journal of Language & Communication Disorders*, *44*(6), 1000–1017. https://doi.org/10.1080/13682820802546951

Sares, A. G., Deroche, M. L. D., Shiller, D. M., & Gracco, V. L. (2019). Adults who stutter and metronome synchronization: Evidence for a nonspeech timing deficit. *Annals of the New York Academy of Sciences*, *1449*(1), 56–69. https://doi.org/10.1111/nyas.14117

SheikhBahaei, S., Millwater, M., & Maguire, G. A. (2023). Stuttering as a spectrum disorder: A hypothesis. *Current Research in Neurobiology*, *5*, 100116. https://doi.org/10.1016/j.crneur.2023.100116

Slis, A., Savariaux, C., Perrier, P., & Garnier, M. (2023). Rhythmic tapping difficulties in adults who stutter: A deficit in beat perception, motor execution, or sensorimotor integration? *PLOS ONE*, *18*(2), e0276691. https://doi.org/10.1371/journal.pone.0276691

Smith, A., & Kleinow, J. (2000). Kinematic correlates of speaking rate changes in stuttering and

normally fluent adults. *Journal of Speech, Language, and Hearing Research : JSLHR*, *43*(2), 521–536.

Team, R. C. (2021). R: A language and environment for statistical computing (Version 4.0. 5)[Computer software]. *R Foundation for Statistical Computing*.

Tilsen, S., & Johnson, K. (2008). Low-frequency Fourier analysis of speech rhythm. *The Journal of the Acoustical Society of America*, *124*(2), EL34–EL39. https://doi.org/10.1121/1.2947626

Toyomura, A., Fujii, T., & Kuriki, S. (2011). Effect of external auditory pacing on the neural activity of stuttering speakers. *NeuroImage*, *57*(4), 1507–1516. https://doi.org/10.1016/j.neuroimage.2011.05.039

White, L., & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*. https://doi.org/10.1016/j.wocn.2007.02.003

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wiltshire, C. E. E., Chiew, M., Chesters, J., Healy, M. P., & Watkins, K. E. (2021). Speech Movement Variability in People Who Stutter: A Vocal Tract Magnetic Resonance Imaging Study. *Journal of Speech, Language, and Hearing Research*, *64*(7), 2438–2452. https://doi.org/10.1044/2021_JSLHR-20-00507

Wiltshire, C. E. E., Cler, G., Chiew, M., Freudenberger, J., Chesters, J., Healy, M., Hoole, P., &

Watkins, K. E. (2023). *Speaking to a metronome reduces kinematic variability in typical speakers and people who stutter*. OSF. https://doi.org/10.31219/osf.io/wc29m

Yairi, E., & Ambrose, N. (2013). Epidemiology of stuttering: 21st century advances. *Journal of Fluency Disorders*, *38*(2), 66–87. https://doi.org/10.1016/j.jfludis.2012.11.002

Zelaznik, H. N., Smith, A., & Franz, E. A. (1994). Motor Performance of Stutterers and Nonstutterers on Timing and Force Control Tasks. *Journal of Motor Behavior*, *26*(4), 340–347. https://doi.org/10.1080/00222895.1994.9941690

Figure 1: Boxplot and connected dot plot showing the rhythm (EMS mean peak frequency) with which neurotypical participants A and B read each passage in the solo and choral conditions. Dots and lines are colour coded by participant pair. Each box represents the interquartile range, with a horizontal line showing the median value. Whiskers extend to 1.5 times the interquartile range.

Figure 2: Boxplot showing the dominant rhythm (EMS mean peak frequency) with which each group read each passage in the solo and choral conditions. Each box represents the interquartile range, with a horizontal line showing the median value. Whiskers extend to 1.5 times the interquartile range. Boxes are colour coded by participant group (green = neurotypical, orange = PWS).

Figure 3. Boxplot showing the consistency of the dominant rhythm (EMS mean peak amplitude) with which each group read each passage in the solo and choral conditions. Each box represents the interquartile range, with a horizontal line showing the median value. Whiskers extend to 1.5 times the interquartile range. Boxes are colour coded by participant group (green = neurotypical, orange = PWS).

Figure 4. Stuttering frequency for each disfluent participant (in percentage syllables stuttered) plotted against their mean EMS peak frequency (averaged across all octave bands, in Hz). Dots are colour coded by condition (red = choral, blue = solo). Shaded area indicates 95% confidence interval around the line of best fit.

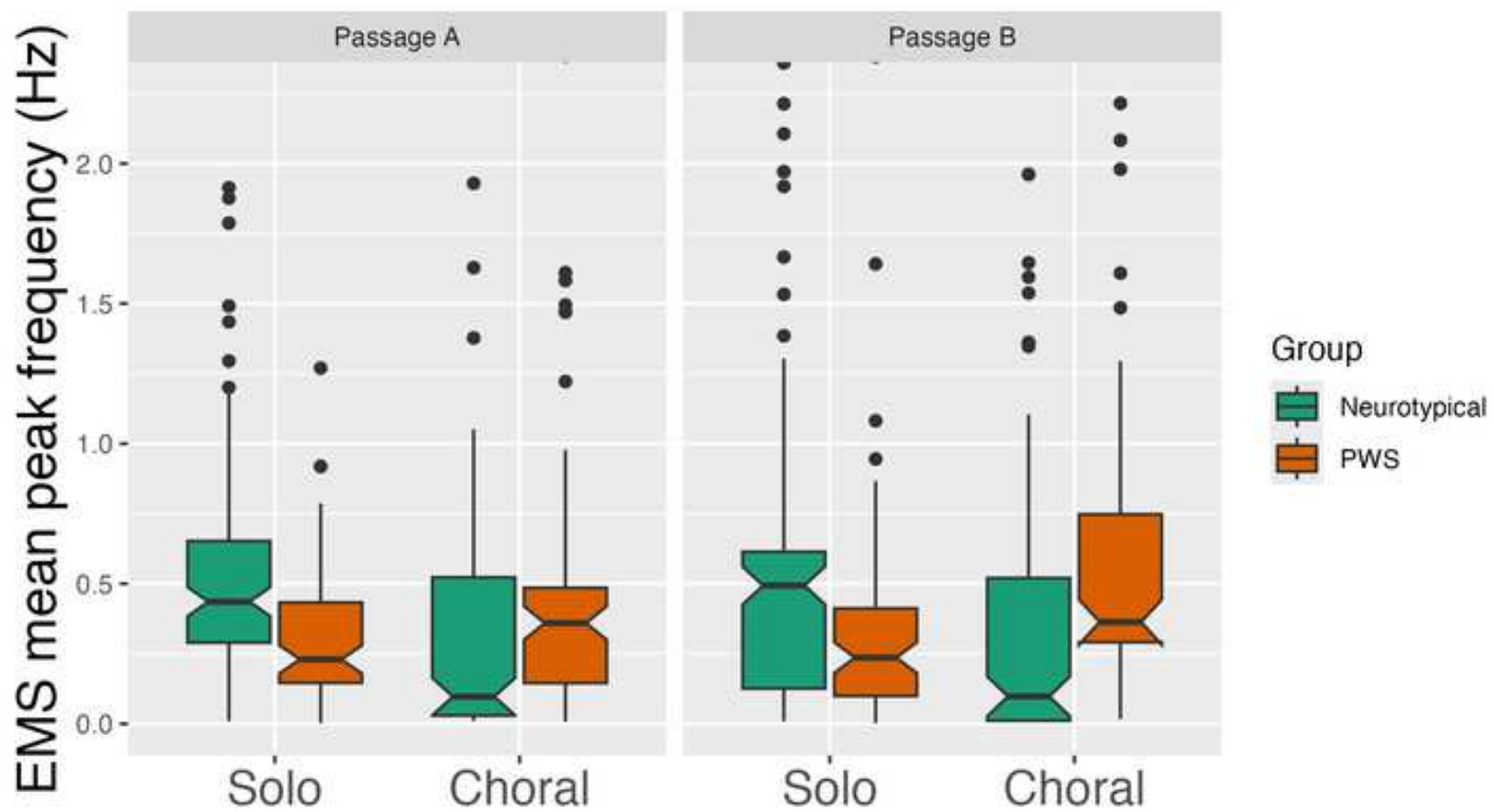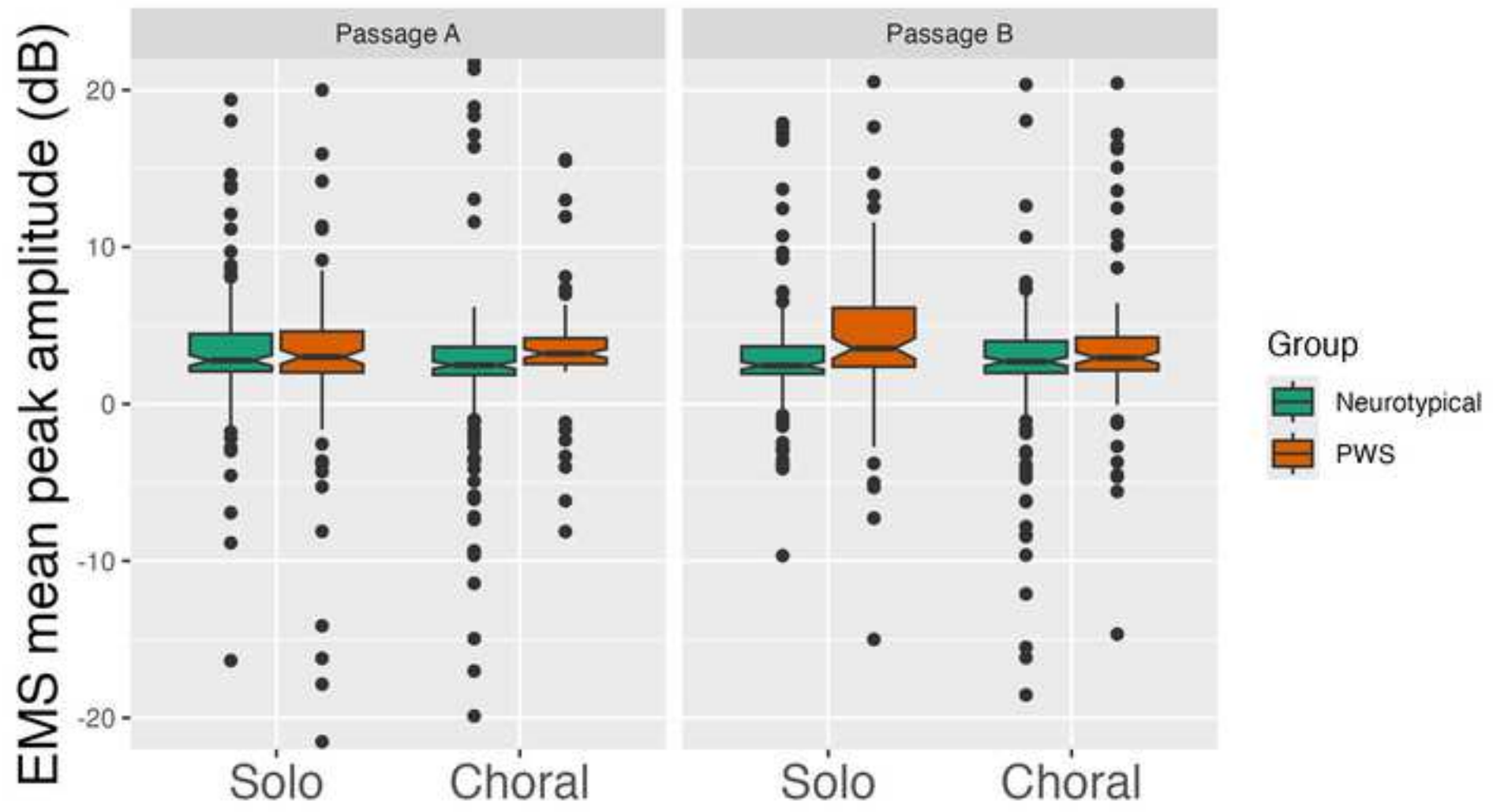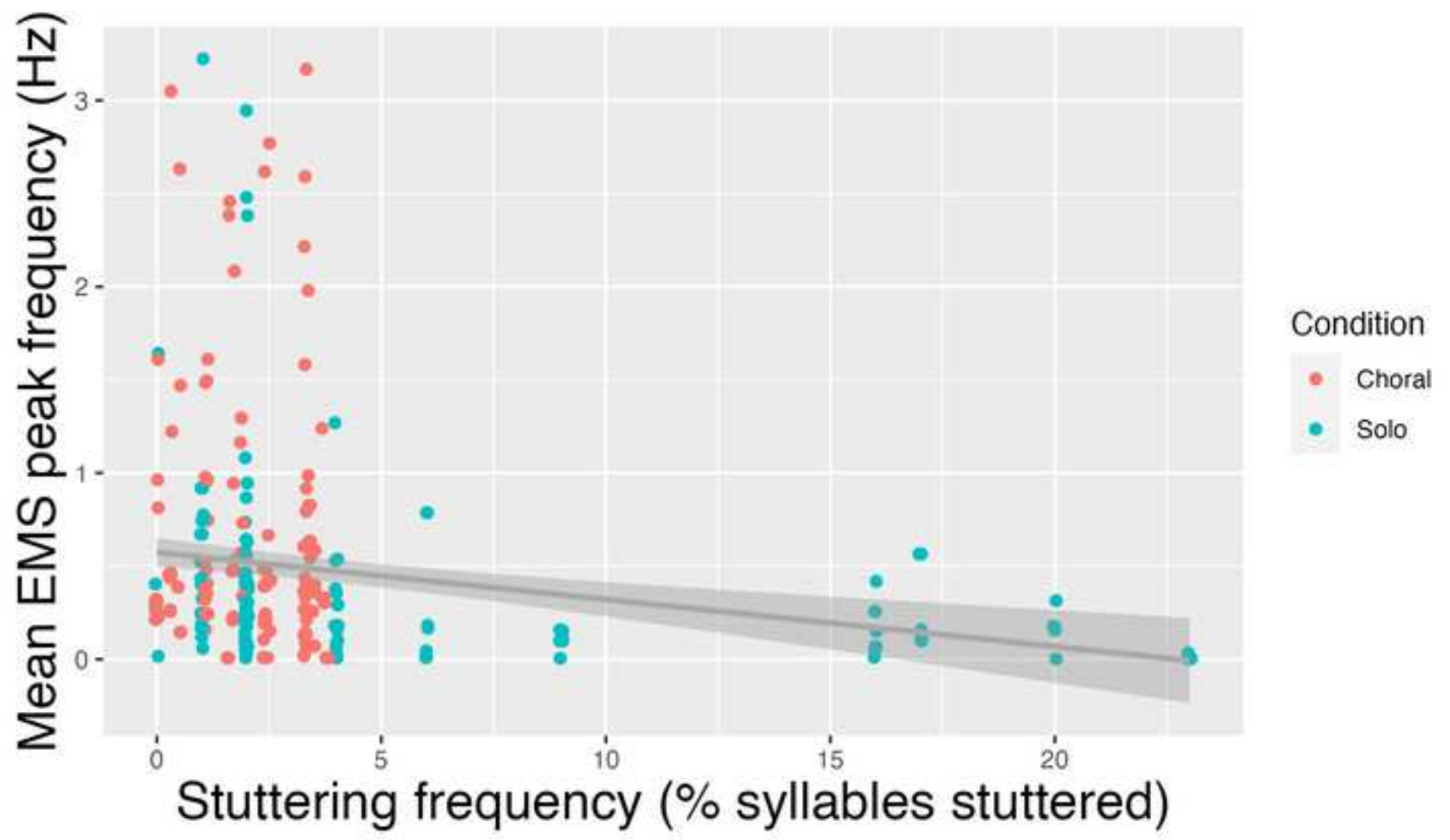Figure 1

Figure 2
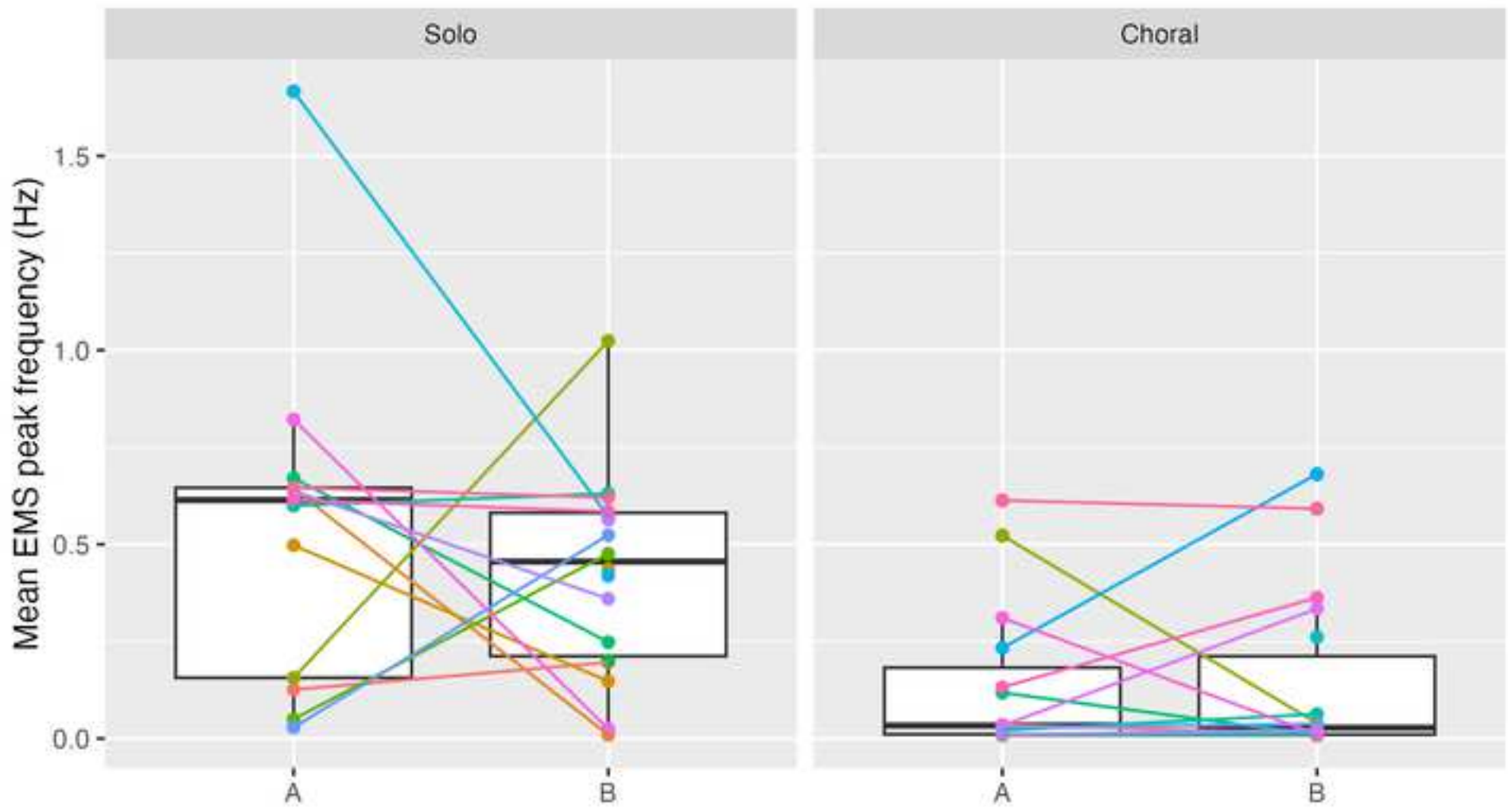
Figure 3

Figure 4

Figure 4

Many thanks to both reviewers for their useful and insightful comments, as well as their kind words on the scientific merit of the paper. In general, we agree with all of the points raised and have incorporated their suggestions into the revised manuscript. A more detailed point-by-point response follows.

Reviewer #1:

Reviewer 1: This is an important paper on how choral speaking increases fluency in stuttering and whether this fluency-enhancing strategy matches with hypothesized motor timing mechanisms. I have some comments to improve the clarity of the paper.

• **Are there any more recent citations on other fluency enhancing strategies other than the citations from the 1950's?**
RESPONSE: We have added some additional citations.

• **I recommend using "people who stutter" and then shortening it to PWS for the rest of the paper. For example, "stuttering talkers" in line 73 is an inappropriate term to use (person-first language is required for this journal) so replacing that with PWS would work well.**
RESPONSE: We have made this change throughout the paper.

• **Lines 76 through 78, I agree with the first two points, but why should speaking chorally change the rhythm to match that of control speakers? This needs to better justified. Line 84 provides a justification but it's not strong enough.**

RESPONSE: We have added the following additional text:

When neurotypical talkers speak chorally, their speech also changes in relatively consistent ways that make their rhythm more predictable, for example by regularising vocalic interval durations (Cerda-Oñate et al., 2021; Cummins, 2009). If people who stutter use their partner's speech as a rhythmic guide during choral reading, then it is to be expected that their speech rhythm will also become more predictable, similar to that of typical talkers during choral speech.

• **Line 89- EMS in choral speaking did not predict fluency during choral speaking or fluency in spontaneous speech? Also, this claim seems to be in contrast to the claim in line 115.**

RESPONSE: We have clarified as follows:

Additionally, Meekings et al. (2023) found that while one measure of rhythm (EMS peak frequency) correlated with induced fluency, the 'rhythmic signature of choral speech' – a combination of EMS metrics that most reliably characterised the choral speech of people who stutter– did not significantly predict fluency in choral speech.

• **The stats and acoustic analyses are well done and robust**

RESPONSE: Thank you!

• Figure captions need to be more detailed to explain what the boxplots mean (dark line, quartile boxes, dots, colors, etc)

RESPONSE: We have done this.

• Can you provide an r squared value between stuttering frequency and EMS peak frequency?

RESPONSE: Yes- now included.

• Line 291 is a really interesting interpretation and makes sense with the results. A follow-up study could look at a choral reading with a recording to read with to see how EMS values change between neurotypical and stuttering group. I think that is what line 293 is saying but it is confusing.

RESPONSE: That is indeed what we meant- have rephrased the line slightly to make this clearer.

Reviewer #2:

Reviewer 2: The study aims to test whether choral speech, a known fluency enhancer for people who stutter, works by giving the speaker an external rhythmic cue to time their speech to. To do this, the study records people who stutter and neurotypical controls as they read a passage with, and without, a choral speech partner. The outcome measure was the peak frequency using an Envelope Modulation Spectrum analysis. Overall, people who stutter had lower peak frequencies (~= faster speech) during solo speech compared with choral speech, whereas controls had higher frequency speech during solo speech compared with the controls, which was then slower during choral speech. The authors suggest that this is evidence that people who stutter do not adjust their speech rhythm to model that of the neurotypical partner during choral speech. This would suggest that people who stutter are not using their partner's rhythm as a way to increase fluency and that another, unknown, mechanism must explain the fluency enhancing effect.

The study is very clearly written with all necessary details to enable replication attempts. The study is well motivated and builds from prior literature. My main concern is whether the research question can be adequately answered based on the design of the study. The authors do note a few of the following concerns in the discussion.

1) The study specifically aims to test a motor timing theory of stuttering. There is a central assumption that people who stutter will move towards the rhythm of the neurotypical speaker because the person who stutters has atypical rhythm and therefore wants to take cues from the neurotypical speaker (lines 83-85). However, during choral speech, both pairs are likely to adapt their speech, which is in accordance with the motor timing hypothesis (i.e. both partners settle on an agreed rhythm early on and then it

becomes consistent and predictable for the person who stutters) and the results of this study (neurotypical speakers adapt their speech during choral reading). However, in the stuttering-neurotypical pairing, the results only show the frequency of the person who stutters. Therefore, it cannot be ruled out that the neurotypical partner was speaking at a higher peak frequency and therefore people who stutter do in fact match the neurotypical speaker. It could be interesting to plot the peak frequency of the individual pairs of neurotypical speakers, to see if they do converge to each other and in which direction, though this still will not address the central question.

RESPONSE: We agree that this is a limitation of the study, as discussed in l.285-294. We have edited that section and moved it to the top of the limitations section to make this clearer, as follows:

Because only one side of the interaction was recorded in the original experiment with PWS, we were unable to compare their speech rhythm directly with that of their partner, meaning that we cannot assess the degree or directionality of convergence in the PWS-experimenter dyads. Instead, we have compared the choral speech behaviour of typical talkers and PWS when paired with a neurotypical partner. Looking at participants' behaviour independent of that of their partner allowed us to control for the possibility that talkers' speech converges during choral reading not because the person with a stutter is converging with the neurotypical talker, but because the neurotypical talker is converging with the person with a stutter. Previous research has found that neurotypical speakers make rhythmic changes during choral speech which make the perceptual centres of their speech more predictable, and that these changes are typically consistent between participants (Cummins, 2009), so we expect neurotypical participants and experimenters to use approximately the same speech rhythm when reading the same passage chorally. However, it may be that typical speakers adopt a different rhythmic strategy when reading chorally with someone who stutters, or that the specific experimenters who partnered with PWS in this study had higher than average peak frequency, meaning that the PWS did in fact entrain to their partner.

We have also plotted neurotypical within-pair convergence and included this in the Results section- thank you for this suggestion!.

2) The issue, in my opinion, is further compounded by differences in the set-up of the choral partners between people who stutter and controls: people who stutter were paired with an experimenter, who likely has knowledge of stuttering and choral reading, and who has likely read the passage many times before and may have spoken the participant before the start of the experiment(?). The control group, on the other hand, were paired with another neurotypical control and were both reading the passage for first time, presumably without any knowledge of choral reading or stuttering. These factors could have quite a significant effect on the results from the perspective of social convergence theory (e.g. Gregory & Webster, 1996 https://psycnet.apa.org/fulltext/1996-01769-009.html) or simple practice effects – see recent paper from Bradshaw & McGettigan for a nice summary (https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0258747)

RESPONSE: This is an interesting and valid point. Research into synchronous speech suggests that experience with choral reading tends not to affect participants' behaviour, and even practising with the same passage only has a marginal effect. However, it is certainly a possibility that the experimenters' experience with the task combined with their higher social status may have driven them to align less with their partner. We have added a paragraph in the discussion relating to this, as follows:

One experimental difference which may have affected the results is the choice of partner. Neurotypical participants were paired with other neurotypical participants, who were of equivalent social status and had limited previous experience of choral speech. Participants with a stutter were paired with one of four experimenters, who had read the passage before and had more experience with choral reading. Previous research has found that, typically, participants are able to read chorally in close synchrony almost immediately, even if they have no prior experience of the task. Moreover, practising a passage only marginally changes choral speech behaviour, while prior experience with choral speech does not affect performance (Cummins, 2003; O'Dell et al., 2010). However, because the experimenter might have been perceived as the dominant partner because they ran the experiment and had experience doing the task, this may have caused the PWS to converge more to the experimenter than vice versa (Gregory & Webster, 1996).

3) **The results, including statistics and figures, are presented clearly. Consideration could be made to a different interpretation: People who stutter have a lower peak frequency during solo reading because they are employing learned or automatic techniques to control stuttering, such as soft onsets, breath control, elongation of initial sounds, or simply slowed speech pace. When they are then in a fluency enhanced condition, they are able to speak with a higher frequency because they no longer need to use techniques. This interpretation would be held whether or not they are using rhythm as a guide or if something else enables them to gain fluency. It also critically depends on the frequency of the experimenter (see point above). The controls on the other hand, do not need to use techniques when speaking in the solo condition, meaning they can speak at a higher frequency compared with people who stutter. When they speak in the choral condition, they slow down to aid the speech matching process. Within-subject plots and/or within-pair plots may help to untangle this.**

RESPONSE: This is an interesting potential explanation, which we have added to the discussion- thank you!

4) **Finally, I understand the rationale for wanting to pick one clear measurement which is based on previous research. However, peak frequency alone, in my opinion, cannot answer the central research question without knowing whether each person in the pair converges towards each other and how consistent the frequency then becomes. If people who stutter match with the choral partner and consistently hit the same rhythm, this would answer the main question. A consistency measure could help to answer**

whether people who stutter maintain a predictable rhythm (even if this information is missing for the choral partner).

RESPONSE: To look at rhythmic consistency, we have included an analysis of peak amplitude, which represents the dominance of the peak frequency. As you'll see, this wasn't super conclusive… but we agree that consistency of speech rhythm could well be an important factor here, and have added a note in the discussion that this should be an avenue of investigation for future studies.

Minor: Can the reasons for the technical difficulties be listed, particularly if this will help people avoid issues in the future?

RESPONSE: This is explained in the Preprocessing section (l. 196-199)- I have additionally amended l. 146 to specify 'issues with recording quality' rather than 'technical issues'.

Minor: Can more details be added from Meekings (2023) justifying why peak frequency is a good measure, what the other measures were and why they were not suitable/chosen for the current project. It is only mentioned that they didn't correlate with stuttering frequency (lines 88, 116).

RESPONSE: We have added the following passage:

Meekings et al. (2023) also identified six EMS metrics that, in combination, significantly predicted whether a participant was speaking chorally or not. These metrics were peak frequency in the 125Hz, 1kHz and 2kHz octave bands, energy between 4-10Hz in the 4kHz band, energy between 3-6Hz in the full band, and ratio of the energy above and below 4Hz in the 8kHz band. Although the weighted linear combination of these variables that predicted choral speech rhythm did not significantly predict fluency, these results do suggest that EMS peak frequency across multiple bands is an important component of choral speech rhythm.