Speaking to a metronome reduces kinematic variability in typical speakers and people who stutter.

Wiltshire, Charlie; Cler, Gabriel J; Chiew, Mark; Freudenberger, Jana; Chesters, Jennifer; Healy, Máiréad P; Hoole, Philip; Watkins, Kate E

PLoS ONE

DOI:
10.1371/journal.pone.0309612

Published: 16/10/2024

Peer reviewed version

Title: Speaking to a metronome reduces kinematic variability in typical speakers and people who stutter.

Authors: Charlotte E.E. Wiltshire[1,2,3], Gabriel J. Cler[1,4], Mark Chiew[1,5,6], Jana Freudenberger[3], Jennifer Chesters[1], Máiréad P. Healy[1,7], Philip Hoole[3], Kate E. Watkins[1].

[1] Wellcome Centre for Integrative Neuroimaging, University of Oxford, UK.
[2] Department of Psychology, School of Human and Behavioural Sciences, Bangor University, UK.
[3] Institute for Phonetics and Speech Processing, Ludwig-Maximilian-University, Munich, Germany.
[4] Department of Speech and Hearing Sciences, University of Washington, USA.
[5] Department of Medical Biophysics, University of Toronto, Toronto, Canada
[6] Physical Sciences, Sunnybrook Research Institute, Toronto, Canada
[7] Department of Psychology, University of Cambridge, UK.

Corresponding Author: Dr. Charlotte Wiltshire
School of Human and Behavioural Sciences,
Bangor University,
LL57 2AS
c.wiltshire@bangor.ac.uk
01248 382211

Abstract

Background: Several studies indicate that people who stutter show greater variability in speech movements than people who do not stutter, even when the speech produced is perceptually fluent. Speaking to the beat of a metronome reliably increases fluency in people who stutter, regardless of the severity of stuttering.

Objectives: Here, we aimed to test whether metronome-timed speech reduces articulatory variability.

Method: We analysed vocal tract MRI data from 24 people who stutter and 16 controls. Participants repeated sentences with and without a metronome. Midsagittal images of the vocal tract from lips to larynx were reconstructed at 33.3 frames per second. Any utterances containing dysfluencies or non-speech movements (e.g. swallowing) were excluded. For each participant, we measured the variability of movements (coefficient of variation) from the alveolar, palatal and velar regions of the vocal tract.

Results: People who stutter had more variability than control speakers when speaking without a metronome, which was then reduced to the same level as controls when speaking with the metronome. The velar region contained more variability than the alveolar and palatal regions, which were similar.

Conclusions: These results replicate previous findings of greater variability in the movements of people who stutter compared with controls during natural, fluent speaking (i.e. speaking without the metronome). Furthermore, these results extend previous knowledge to show that in addition to increasing fluency in people who stutter, metronome timed speech also reduces articulatory variability to the same level as that seen in control speakers.

Introduction

Developmental stuttering is characterised by unplanned interruptions to the flow of speech. Perceivable characteristics can include tense pauses, repetitions, and prolongations of sounds. These characteristics vary greatly between people and day-to-day within an individual. Despite the heterogeneity of stuttering, there are numerous 'fluency-enhancing' techniques that reliably reduce the likelihood of stuttering. Using these techniques can result in fluent speech for people who stutter regardless of the day-to-day amount of stuttering for the individual speaker. These fluency enhancers often use an external timing cue to aid speech, including metronome timed speech, speaking in unison with another speaker, and singing.

This 'rhythm effect' can be explained at a neurobiological level by two neural loops that pass through the basal ganglia [1]: an "external timing network" comprising the basal ganglia and supplementary motor area (SMA) and an "internal timing network" made up of the premotor cortex, cerebellum and basal ganglia. Evidence for the dissociation between internal and external timing networks for movement comes from animal and patient studies. Recordings from neurons in macaque SMA showed a preference for internally generated movements (producing a predetermined sequence without a cue) whereas neurons in premotor cortex showed a preference for movement sequences that were cued externally (via lights) [2]. Patients with lesions involving the SMA and basal ganglia often have difficulties when producing internally generated movements but benefit substantially when moving or speaking with an external cue [3–5]. Similarly, people with Parkinson's Disease have dysfunction in the basal ganglia motor loops and often have difficulties with limb (and sometimes speech) movements that are internally generated, but less so during externally cued movements [6–10].

Neurobiological explanations of stuttering propose that signalling within the SMA-basal ganglia loop is noisy, which leads to breakdown of the speech stream [1,11]. The cerebellum-premotor-basal ganglia loop, on the other hand, may act to support fluent speech production when speaking with a predictable external cue [1]. Evidence from structural and functional brain imaging studies show differences in the basal ganglia in people who stutter [12,13]. The amount of activity within the substantia nigra during fluent speech has been shown to correlate with clinical measures of stuttering "severity" in people who stutter [14–16]. Structurally, the left putamen and left cortical speech areas were found to have higher iron levels in people who stutter, which may be indicative of elevated dopamine levels [17]. Dopamine is critically involved in regulating timing cues and the balance between inhibitory and excitatory signals in the basal ganglia-cortical networks. Together, there is consistent evidence that the basal ganglia motor loops are important in the cause of stuttering, and that there is a dissociation between how these loops control internally generated and externally cued (speech) movements.

Research into the neural basis of the "rhythm effect" in people who stutter is limited. However, in a recent study, fMRI was used to compare blood-oxygen-level-dependent (BOLD) responses when people who stutter read a passage with, and without, an imagined metronome [18]. The stuttering group had increased functional connectivity involving the cerebellum during

externally paced reading compared with reading without an imagined metronome. The control group, on the other hand, did not show additional recruitment of the cerebellum in the metronome condition and instead had greater activity in the SMA. This suggests that people who stutter recruit the cerebellum during a metronome-timed speech condition as an "organic attempt at compensation" [18]. This lends support to the idea fluency enhancement is achieved because of the engagement in the cerebellum-premotor-basal ganglia loop when an external cue, such as a metronome, is present [1].

The behavioural effects of metronome-timed speech in enhancing fluency in people who stutter are well established. However, these rely on acoustic or perceptual measures of speech and speech fluency. Previous studies have found that during natural speech, the variability of speech movements, that is, how similar the speech movements are to each other over repeated utterances, is higher in the stuttering group compared with the control group. Higher articulatory variability in fluent speech has been reported in studies that record repeated utterances [19], vowel sounds [20], targeted jaw movements [21], simple and complex pseudowords [19,22,23] and simple sentences [24–26].

The finding of higher trait-level variability in articulatory movements in people who stutter, though consistent, is difficult to interpret. The prevailing interpretation is that greater variability could indicate 'weaker' speech motor control at the trait level. This weakness could then contribute to the breakdown of speech motor control leading to a stutter [19]. Importantly, none of the prior studies report a relationship between clinical measures of stuttering (e.g. the SSI-4 [27]) and the trait-level amounts of variability. This suggests that higher variability is not the predominant cause of stuttered speech.

Therefore, while trait-level variability appears to be an important difference in people who stutter, the reason for higher variability and the relationship between variability and stuttered moments is unknown. The remaining questions are whether variability reduces when demands on the speech motor system also reduce and whether "fluency enhancing" techniques work to reduce variability in addition to reducing perceptible stuttering moments.

Here, we first aimed to reproduce previous findings that people who stutter have greater variability in their speech movements compared with people who do not stutter. Further, we aimed to extend our knowledge by investigating the effect of metronome-timed speech on speech movement variability. A reduction in variability during metronome-timed speech would indicate that, given the right conditions, there is nothing fundamentally different in the capabilities of the speech motor control system in people who stutter compared with control speakers. It would also indicate that speaking with a fluency enhancer directly leads to a change in the speech motor control system. We leverage the advantages of vocal tract MRI to investigate speech movements along the length of the vocal tract; the lips, tongue and velum.

Method

Participants

We scanned 27 people who stutter (PWS) and 20 control participants (CON). Data were collected from the same participants and session as our previous study [23], though four participants who stutter did not take part in the current task due to time constraints or

technical difficulties. Data sets from five participants (2 PWS, 3 CON) were excluded due to technical difficulties with the scanner and audio set up. One dataset (PWS) was removed as there was an insufficient number of fluent utterances for analysis purposes (see analysis plan section) and one dataset (CON) was excluded due to the words being pronounced in different ways over the repetitions. This resulted in a sample of 24 adults who stutter and 16 controls. The groups were balanced for gender, age and years of education (Table S1). All people who stutter had at least a score of 18, corresponding to "mild" on the Stuttering Severity Instrument-4 [27]. Participants reported normal or corrected-to-normal vision and normal hearing. No participants had taken part in speech therapy within the last six months. During the experiment, participants were asked not to actively use any of techniques that might help them to maintain fluency.

The University of Oxford Central University Research Ethics Committee (R52173/RE005) approved the study. Participants gave informed written consent to participate in the study, in accordance with the Declaration of Helsinki, and with the procedure approved by the committee.

Experimental Procedure

Prior to the experiment, the participants practised speaking the stimuli out loud. During scanning, for both the metronome and natural condition, three sentences were repeated 10 times each in a random order (total 60 trials). For each trial, the sentence was displayed on the screen and participants read the sentence out loud. For the metronome condition ("Met"), participants heard a metronome (2Hz). For the no-metronome condition ("NoMet"), participants spoke at their natural speaking rate. Each trial lasted seven seconds. Each sentence contained a multisyllabic target word from Motor Speech Evaluation [28] (see Analysis Procedure).

MRI Acquisition

Data were obtained using a 3-T MRI system (Prisma, Siemens) with a 64-channel head and neck receive array. Midsagittal images of the vocal tract were acquired with in-plane spatial resolution of 2mm X 2mm using a radial FLASH sequence (echo time/repetition time – 1.4/2.5ms) with golden-angle sampling. Images were reconstructed at 33.3 frames per second using a second-order spatiotemporal total generalized variation constraint [29].
Audio was recorded using noise-cancelling procedures (Optoacoustics, FOMRIII).

Analysis Procedure

The audio and reconstructed videos of the vocal tract were synchronized and inspected for any disfluencies or non-speech movements including yawning and swallowing. If these movements affected the production any part of the sentence, the trial was excluded. The audio was then segmented acoustically in Praat [30]. The target utterance was extracted from the sentence based on clear acoustic signals before and after the target word. These were: 1) the release of /ð/ and closure formation of /v/ in "Bob knew th*e impossibility o*f the task" 2) release of /d/ and closure formation of /p/ in "Mike carrie*d artillery u*p the hill" 3) release of /d/ and closure formation of /t/ in "Peter ha*d a catastrophe a*t work".

The vocal tract images were segmented based on the manual acoustic segmentation. Vocal tract data were analysed using an open-source vocal tract MRI toolbox [31]. Whilst this toolbox

is different from that used in our first vocal tract study [23], the analysis procedures are very similar. First, a frame with an open airway is chosen. Then, the glottis, velopharyngeal port, and alveolar ridge are marked manually. The midpoint of the line from the alveolar ridge to the glottis was located within the genioglossus muscle. This served as the origin for a semi-polar grid with 28 gridlines. The gridlines terminated at the automatically detected posterior or superior boundary. The mean pixel intensity is calculated for each gridline. The grid was then further divided into articulatory regions along the length of the vocal tract (See Figure 1). Therefore, the signal from each region is the mean of the mean pixel intensity for the relevant gridlines, scaled between zero and one for each participant. A large number of high-intensity pixels (i.e. tissue) along the gridline indicates high constriction of the vocal tract region. For the current analysis, we focussed on the alveolar, palatal and velar regions.

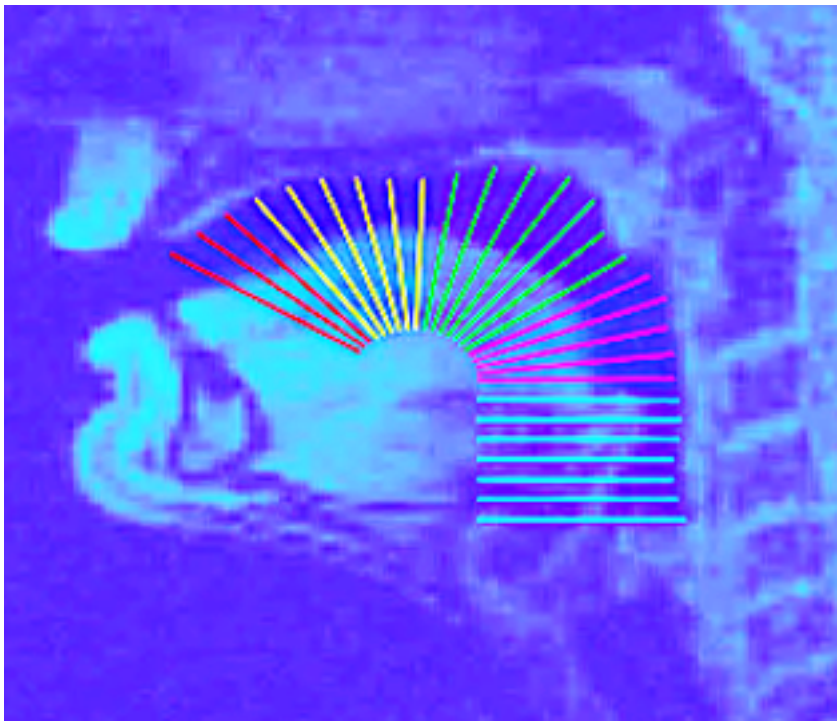

Figure 1: A frame with an open vocal tract (in this case, the schwa vowel in "the hill" during the metronome condition). Dark blue is air and bone, light blue is tissue. A semi-polar grid is constructed over the airway (see analysis procedures for details). Gridlines were coloured by region. Red = Alveolar, yellow = Palatal, green = Velar, pink = Hyperpharyngeal, blue = Hypopharyngeal.

Measured Variables

Variability was calculated using the coefficient of variation (CoV), that is, the standard deviation of the size of the movements across 10 repetitions of each target utterance, divided by the mean. The size of the movement was simply the sum of the constriction value of the movements across frames capturing both the amplitude and duration of the movement (see Analysis Procedure and Figure 2).

Movement duration was recorded for each repetition by summing the total number of frames from the start to the end of the target utterance as defined above.

Analysis Plan

Trials resulting in exclusion were rare. If a participant did not produce at least six (out of 10) fluent and accurate productions of a sentence, data for that sentence were excluded from analyses. Trials containing stuttering were excluded here but will be used in future work [32]. Prior to analysis, two full data sets were removed; one person who stutters due to stuttering and one control participant due to mispronunciations. After removing these participants, a further 4.2% of sentences from the control group and 2% of sentences from the stuttering group were excluded (Table S2). These exclusions were considered missing at random.

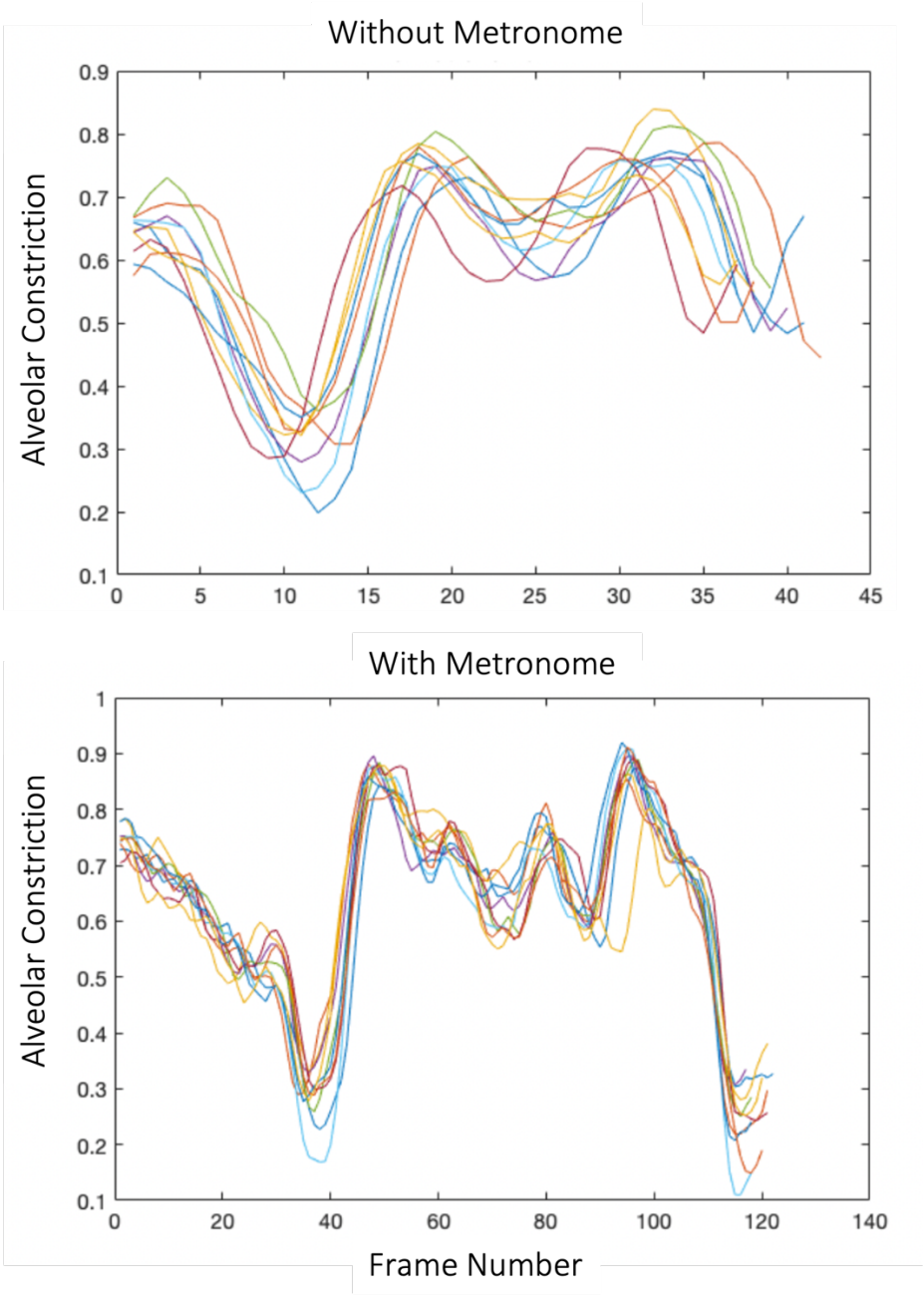

Figure 2. Example alveolar movement traces from one participant who stutters. Ten repetitions of the target utterance "*the impossibility of*" are overlaid (see "Analysis Procedure", above). Each line represents one repetition. Data were reconstructed at 33.3 frames per second. Low constriction values mean the airway was more open.

Results

Summary statistics are given in Table S2. The amount of variability (CoV) for each condition (with or without metronome) and each region (alveolar, palatal and velar) is plotted in Figure 3. A summary variability score for each participant (collapsed across region) is plotted in Figure 4. Duration of the utterances for condition ("Met", "NoMet") are plotted in Figure 5.

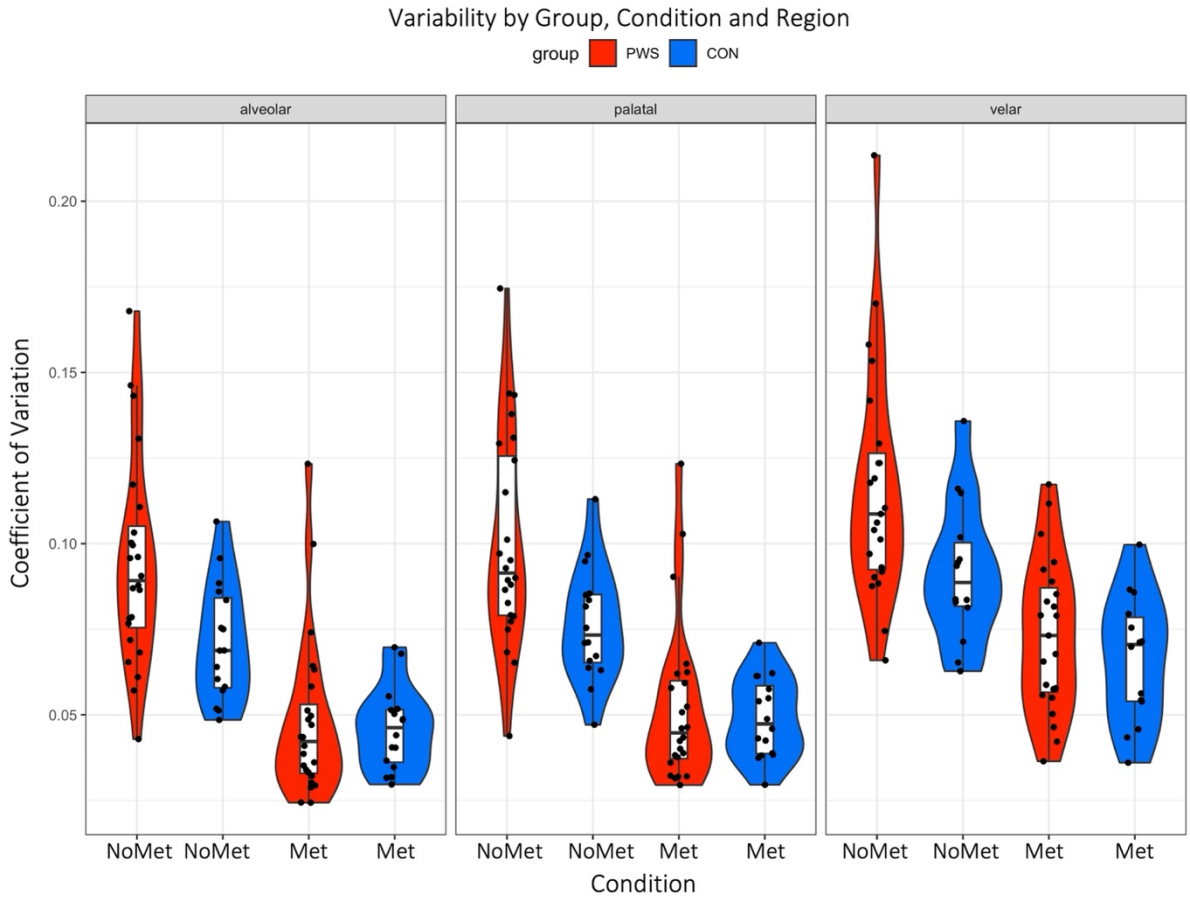

Figure 3. Coefficient of variation by group (PWS, CON), condition ("Met" and "NoMet") and region (alveolar, palatal, velar). Each point is the variability of one participant averaged across the three sentences. Violin plots are shown to visualize the distribution of data and its probability density for each group and condition separately for each region of the vocal tract. Boxes display median and interquartile range; whiskers represent an additional 1.5*IQR.

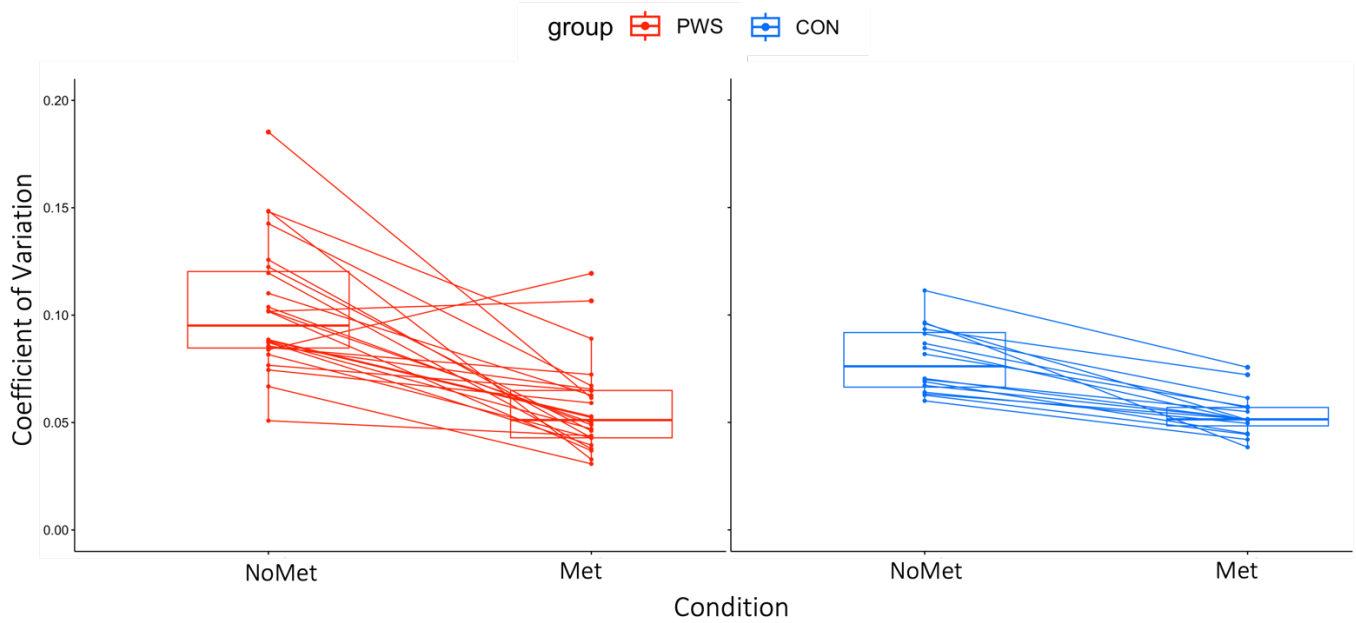

Figure 4. Interaction between condition ("Met" and "No Met") and group. A line connects each participants data between the conditions. Data are averaged for each participant over region (alveolar, palatal, velar) and sentence. Boxplots display median and IQR, whiskers represent an additional 1.5*IQR.

We examined whether variability in speech movements during fluent repetitions of simple sentences differed 1) between people who stutter and a control group and 2) when speaking with, or without, a metronome. To do this, the *brms* package [33] was used to fit two Bayesian regression models. In the first model, we tested whether the coefficient of variation during the "NoMet" condition (i.e. natural speech) could be predicted by group (people who stutter vs. the control group) and region (alveolar, palatal and velar). In the second model, we tested whether coefficient of variation could be predicted by group, condition and region, as well as the interaction between group and condition. Region was not expected to interact with either of the other variables and therefore was not included as an interaction term. For both models, sentence was nested within participant a random factor. Based on previous literature [23], mildly informative priors for the coefficient of variation were set between zero and 0.5. A lognormal data distribution gave the best fit to the models. These models fitted well with *rhats* uniformly at 1, a large number of effective samples as well as stationary and well mixing chains (see R code for details of diagnostics).

Differences in variability between groups

Variability (coefficient of variation) was predicted by group and region (Figure 3) using the first model. The stuttering group spoke with higher variability than the control group (β=0.19; lower 95% credible interval (CI)=0.02, upper 95% credible interval (CI)=0.37). The velar region moved with more variability than both the alveolar (β=0.24; CI=0.19-0.21) and palatal regions (β=0.17; CI=0.12-0.22), which were similar (β=0.07; CI=0.02-0.12).

The 'Metronome Effect' on Variability

In the second model, the coefficient of variation was predicted by condition, with higher variability for the "NoMet" compared with the "Met" condition (β=0.42; CI=0.34-0.51) and an interaction between group and condition, with the stuttering group showing a larger reduction in variability from the "NoMet" to "Met" conditions compared with controls (β=0.20; CI=0.09-0.31). As with the first model, velar movements were more variable compared with both alveolar and palatal movements, which were similar.

Relationship between stuttering and variability

We selected the coefficient of variation for the "NoMet" condition, averaged across sentences and regions, as an overall measure of variability. There was no correlation between variability score and SSI-4 score (r=0.21, p=0.33).

Duration

The duration of the "Met" condition was considerably longer than the "NoMet" condition, as expected, given the pace of the metronome (2Hz). There were no differences in average duration or variance of the durations between the groups for the "NoMet" or Met conditions. As expected, given the predictable timing constraint of the metronome, the variance of the durations was larger for the "NoMet" condition compared with the Met condition for both groups.

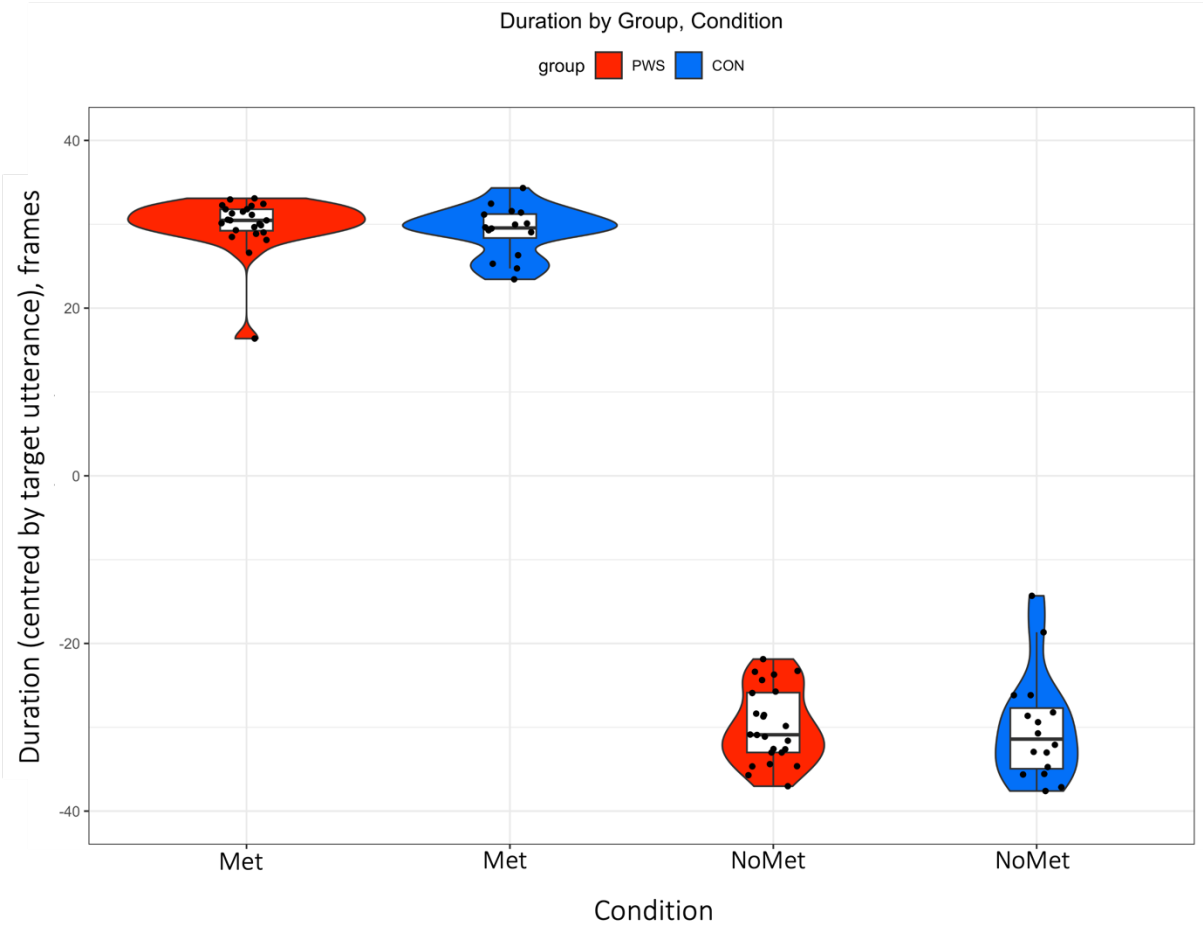

Figure 5. Average duration for each participant for the metronome ("Met") and no metronome ("NoMet") condition measured in frames (images reconstructed at 33 frames per second). The mean of duration was centred for each of the three target utterances prior to averaging to account for differences in utterance length between the three utterances. Violin plots are shown to visualize the distribution of data and its probability density for each group and condition separately. Boxplots display median and interquartile range; whiskers represent an additional 1.5*IQR.

Discussion

Summary of findings

We used vocal tract MRI to capture movements within the alveolar, palatal and velar regions at high temporal and spatial resolution in 24 people who stutter and 16 controls. We recorded images of the vocal tract as participants repeated simple sentences with and without a known "fluency-enhancer", metronome-timed speech. First, we report that the stuttering group spoke with greater variability compared with the control group during natural speech that was perceived as fluent (i.e. without the metronome). Secondly, we show that both people who stutter and control speakers reduce their speech motor variability when speaking to a metronome, and that this effect is larger in the stuttering group. Overall, the variability of the stuttering group reduced to the same level as the control group when speaking to a metronome. Finally, we found no correlation between kinematic variability and clinical measures of stuttering "severity".

Greater variability in the fluent speech of people who stutter

The current work adds to a substantial body of work demonstrating that during speech that is perceived as fluent, people who stutter speak with higher amounts of articulator variability compared with control participants. This result, therefore, represents a robust finding in the literature that can be reproduced using different methods (e.g. Optotrak, vocal tract MRI), stimuli (sounds, nonwords, sentences) and analysis procedures (including coefficient of variation, as used here, and the spatiotemporal index [34]). It is unclear, however, why the stuttering group speak with higher variability and what the underlying mechanism is. To understand the factors that influence speech motor variability, we added the metronome condition to our experimental design, to understand whether variability can be altered by a known "fluency enhancer".

Speaking to a metronome reduces variability

During the metronome condition, there were no differences in variability between the groups. These results suggest that, given supportive conditions for fluent speech (i.e. with a "fluency enhancer"), people who stutter show no differences in speech motor control compared with people who do not stutter. This shows that there is nothing fundamentally restrictive about the speech motor control system in developmental stuttering.

Metronome-timed speech is thought to work by providing the speaker with an external, predictable timing cue. Our results show that both controls and people who stutter reduced the variability in their speech when speaking to the beat of a metronome. This is perhaps not surprising, given that our measure of variability, the coefficient of variation, captures variance

in the duration of the utterance and that the metronome beat restricts the duration of the utterance. Instead, the effect of interest is that the stuttering group reduced the level of variability to that of controls when speaking to a metronome.

No relationship between variability and stuttering "severity"

Our results clearly demonstrate that there was no relationship between a clinical measure of stuttering (SSI-4) and the amount of variability in fluent speech for the stuttering group. Furthermore, not everyone in the stuttering group has variability values outside of the range of the control group during the natural speech (no-metronome) condition. This replicates previous findings [19,23] and suggests it is unlikely that the main mechanism behind stuttering is high levels of variability that lead to a breakdown of the speech stream, resulting in moments of stuttering. With this in mind, we discuss two opposing explanations for the role of kinematic variability in developmental stuttering.

What does "more variability" mean for people who stutter?

The prevailing interpretation is that greater variability could indicate 'weaker' speech motor control at the trait level. That is, signalling between key speech motor control regions of the brain is disrupted in the brains of people who stutter and this disruption leads to instability in the kinematic output. This has been suggested to contribute to the disruption of speech motor control leading to a stutter [19]. An alternative interpretation is that greater variability could be the result of attempts to maintain fluency. That is, having a more flexible and adaptive speech motor control system may allow people who stutter to avoid stuttering. Our results could provide support for both options. In both scenarios, the demands on the speech motor control system are higher when speaking without a "fluency enhancer". These increased demands lead to higher variability due either to weaker speech motor control or to an increased need for a flexible and adaptive system. Again, in both scenarios, demands on the speech motor system are reduced when the external cue acts as a scaffold, resulting in less variability.

Variability throughout the vocal tract

We divided the vocal tract into three sections; the alveolar, the palatal and velar regions. The alveolar and palatal regions produced similar results to each other in both groups of speakers. The velar region, however, had higher variability scores compared with the alveolar and palatal region in both groups. This result is in the opposite direction of a previous study that used vocal tract MRI, which reported lower velum variability compared with alveolar and palatal regions [23]. The reason for higher variability in the velar region is unclear. As with the prior study, our three target utterances did not require much movement within the velar region and there are very few sounds with strong constraints on the position of the tongue in the velar region. One explanation is that the temporal resolution is not fast enough to reliably capture the change in the velum moving from a closed to an open position. This particular result requires more investigation. Vocal tract MRI is a rapidly evolving field, and it is now possible to record images at over 100 fps [35]. Higher temporal and spatial resolution may help to resolve these conflicting results.

Limitations

One key limitation of this work is that the metronome was slower than the participants' natural speech rate and therefore it is difficult to separate the effects of the metronome and the effects of simply slowing speech. Stuttering frequency reduces at slower speech rates [36,37]. In addition, our previous findings [23] showed that people who stutter reduced the speed of their speech when productions were more phonologically complex. Taken together, this suggests that slowing speech is a useful mechanism for maintaining fluency for people who stutter and may contribute to the reduction in variability when speaking in the metronome condition.

The SSI is a tool that is commonly used to measure the severity of stuttering in individuals who stutter. However, the SSI only measures the severity of stuttering based on the frequency of stuttering events and the amount of visible tension there is during speech. Furthermore, it also only captures speech at one point in time which is particularly important given the high degree of day to day or moment to moment fluctuation that can occur in the fluency of speech for people who stutter. The lack of relationship between SSI score and variability score is could be due to problems with the measurement tool. However, we believe this is unlikely given that previous work has also failed to find a correlation. It is more likely, in our opinion, that clinical measures of stuttering are not linked to speech motor variability.

Future Directions

Other populations of speakers, including people with Parkinson's Disease, SMA syndrome and Tourette's Syndrome also show fluency enhancing effects from external cueing techniques. Vocal tract MRI could be used with these other speaker populations to measure motor control or track therapeutic outcomes. We have demonstrated that speech motor variability can be captured non-invasively and quickly (10 minutes), making this approach suitable for these populations. Variability is also routinely measured in the general movement domain to look at whole body or limb movements. The resulting kinematic data is very similar to that produced by vocal tract MRI and as such, these multiple movement modalities could be compared directly, within participants, to understand the motor system as a whole.

Acknowledgements

Charlotte Wiltshire was supported by a DPhil scholarship from the Economic and Social Research Council UK (ES/J500112/1) and the Engineering and Physical Science Research Council UK (EP/N509711/1) and a postdoctoral fellowship from Ludwig-Maximilians-University, Munich. This fellowship received funding from the European Union's Framework programme for Research and Innovation Horizon 2020 under the Marie Skłodowska-Curie Grant Agreement No. 754388 (LMU Research Fellows) and from LMUexcellent, funded by the Federal Ministry of Education and Research (BMBF) and the Free State of Bavaria under the Excellence Strategy of the German Federal Government and the Länder. Mark Chiew was supported by the Royal Academy of Engineering (RF201617\16\23). This work was supported by the Medical Research Council UK Grant MR/N025539/1. The Wellcome Centre for Integrative Neuroimaging was supported by core funding from the Wellcome Trust (203139/Z/16/Z). The authors would like to thank Juliet Semple, Nicola Aikin, Nicola Filippini, and Stuart Clare for their MRI support, Louisa Needham for her assistance with recruitment, Magdalena Saumweber for her assistance with data processing, and Sam Jones for useful feedback on the statistical analysis. We would like to thank Aivy Nguyen, Timothy Berezhnoy, and Anna Nolan for their assistance in acoustic segmentation. Lastly, but certainly not least, the authors would like to thank all the participants who took part in this study.

References

[1] Alm PA. Stuttering and the basal ganglia circuits: a critical review of possible relations. J Commun Disord 2004;37:325–69. https://doi.org/10.1016/j.jcomdis.2004.03.001.

[2] Mushiake H, Inase M, Tanji J. Neuronal activity in the primate premotor, supplementary, and precentral motor cortex during visually guided and internally determined sequential movements. J Neurophysiol 1991;66:705–18. https://doi.org/10.1152/jn.1991.66.3.705.

[3] Ackermann H, Ziegler W. Brain Mechanisms Underlying Speech Motor Control. The Handbook of Phonetic Sciences, Wiley; 2010, p. 202–50. https://doi.org/10.1002/9781444317251.ch6.

[4] Ziegler W, Kilian B, Deger K. The role of the left mesial frontal cortex in fluent speech: Evidence from a case of left supplementary motor area hemorrhage. Neuropsychologia 1997;35:1197–208. https://doi.org/10.1016/S0028-3932(97)00040-7.

[5] Ackermann H, Hertrich I, Ziegler W, Bitzer M, Bien S. Acquired dysfluencies following infarction of the left mesiofrontal cortex. Aphasiology 1996;10:409–17. https://doi.org/10.1080/02687039608248420.

[6] Goberman AM, Blomgren M, Metzger E. Characteristics of speech disfluency in Parkinson disease. J Neurolinguistics 2010;23:470–8. https://doi.org/10.1016/J.JNEUROLING.2008.11.001.

[7] Skodda S, Schlegel U. Speech rate and rhythm in Parkinson's disease. Movement Disorders 2008;23:985–92. https://doi.org/10.1002/mds.21996.

[8] Siegert RJ, Harper DN, Cameron FB, Abernethy D. Journal of Clinical and Experimental Neuropsychology Self-Initiated Versus Externally Cued Reaction Times in Parkinson's Disease 2010. https://doi.org/10.1076/jcen.24.2.146.991.

[9] Jacobs J v., Lou JS, Kraakevik JA, Horak FB. The supplementary motor area contributes to the timing of the anticipatory postural adjustment during step initiation in participants with and without Parkinson's disease. Neuroscience 2009;164:877–85. https://doi.org/10.1016/j.neuroscience.2009.08.002.
[10] Späth M, Aichert I, Timmann D, Ceballos-Baumann AO, Wagner-Sonntag E, Ziegler W. The role of the basal ganglia and cerebellum in adaptation to others' speech rate and rhythm: A study of patients with Parkinson's disease and cerebellar degeneration. Cortex 2022;157:81–98. https://doi.org/10.1016/J.CORTEX.2022.08.012.
[11] Busan P. Developmental stuttering and the role of the supplementary motor cortex. J Fluency Disord 2020;64:105763. https://doi.org/10.1016/j.jfludis.2020.105763.
[12] Tourville JA, Guenther FH. The DIVA model: A neural theory of speech acquisition and production. Lang Cogn Process 2011;26:952–81. https://doi.org/10.1080/01690960903498424.
[13] Chang S-E, Guenther FH. Involvement of the Cortico-Basal Ganglia-Thalamocortical Loop in Developmental Stuttering . Frontiers in Psychology 2020;10.
[14] Giraud AL, Neumann K, Bachoud-Levi AC, von Gudenberg AW, Euler HA, Lanfermann H, et al. Severity of dysfluency correlates with basal ganglia activity in persistent developmental stuttering. Brain Lang 2008;104:190–9. https://doi.org/10.1016/j.bandl.2007.04.005.
[15] Watkins KE, Smith S, Davis S, Howell P. Structural and functional abnormalities of the motor system in developmental stuttering. Brain 2008;131:50–9. https://doi.org/10.1093/brain/awm241.
[16] Metzger FL, Auer T, Helms G, Paulus W, Frahm J, Sommer M, et al. Shifted dynamic interactions between subcortical nuclei and inferior frontal gyri during response preparation in persistent developmental stuttering. Brain Struct Funct 2018;223:165–82. https://doi.org/10.1007/s00429-017-1476-1.
[17] Cler GJ, Krishnan S, Papp D, Wiltshire CEE, Chesters J, Watkins KE. Elevated iron concentration in putamen and cortical speech motor network in developmental stuttering. Brain 2021;144:2979–84. https://doi.org/10.1093/brain/awab348.
[18] Frankford SA, Heller Murray ES, Masapollo M, Cai S, Tourville JA, Nieto-Castañón A, et al. The Neural Circuitry Underlying the "Rhythm Effect" in Stuttering. Journal of Speech, Language, and Hearing Research 2021;64:2325–46. https://doi.org/10.1044/2021_JSLHR-20-00328.
[19] Smith A, Sadagopan N, Walsh B, Weber-Fox C. Increasing phonological complexity reveals heightened instability in inter-articulatory coordination in adults who stutter. J Fluency Disord 2010;35:1–18. https://doi.org/10.1016/j.jfludis.2009.12.001.
[20] Frisch SA, Maxfield N, Belmont A. Anticipatory coarticulation and stability of speech in typically fluent speakers and people who stutter. Clin Linguist Phon 2016;30:277–91. https://doi.org/10.3109/02699206.2015.1137632.
[21] Loucks TM, de Nil L. Oral kinesthetic deficit in adults who stutter: A target-accuracy study. J Mot Behav 2006;38:238–46. https://doi.org/10.3200/JMBR.38.3.238-247.
[22] Sasisekaran J. Nonword repetition and nonword reading abilities in adults who do and do not stutter. J Fluency Disord 2013;38:275–89. https://doi.org/10.1016/j.jfludis.2013.06.001.
[23] Wiltshire CEE, Chiew M, Chesters J, Healy M, Watkins KE. Speech movement variability in people who stutter: A vocal tract MRI study. Journal of Speech, Language, and Hearing Research 2021.
[24] Howell P, Anderson AJ, Bartrip J, Bailey E. Comparison of acoustic and kinematic approaches to measuring utterance-level speech variability. Journal of Speech,

Language, and Hearing Research 2009;52:1088–96. https://doi.org/10.1044/1092-4388(2009/07-0167).
[25] Jackson ES, Tiede M, Beal D, Whalen DH. The impact of social–cognitive stress on speech variability, determinism, and stability in adults who do and do not stutter. Journal of Speech, Language, and Hearing Research 2016;59:1295–314. https://doi.org/10.1044/2016_JSLHR-S-16-0145.
[26] MacPherson MK, Smith A. Influences of sentence length and syntactic complexity on the speech motor control of children who stutter. Journal of Speech, Language, and Hearing Research 2013;56:89–102. https://doi.org/10.1044/1092-4388(2012/11-0184).
[27] Riley G. SSI-4 stuttering severity instrument fourth edition 2009.
[28] Wertz RT, LaPointe LL, Rosenbek JC. Apraxia of speech in adults: The disorder and its management. Singular Publishing Group; 1991.
[29] Knoll F, Bredies K, Pock T, Stollberger R. Second order total generalized variation (TGV) for MRI. Magn Reson Med 2011;65:480–91. https://doi.org/10.1002/mrm.22595.
[30] Boersma P, Weenink D. Praat: doing phonetics by computer . Http://WwwPraatOrg/ Version 6223 2022.
[31] Carignan C. MRI-analyses 2019.
[32] Lu Y, Wiltshire CEE, Watkins KE, Chiew M, Goldstein L. Characteristics of articulatory gestures in stuttered speech: A case study using real-time magnetic resonance imaging. J Commun Disord 2022;97. https://doi.org/10.1016/J.JCOMDIS.2022.106213.
[33] Bürkner PC. brms: An R package for Bayesian multilevel models using Stan. J Stat Softw 2017;80. https://doi.org/10.18637/JSS.V080.I01.
[34] Smith A, Goffman L, Zelaznik HN, Ying G, McGillem C. Spatiotemporal stability and patterning of speech movement sequences. Exp Brain Res 1995;104:493–501. https://doi.org/10.1007/BF00231983.
[35] Lingala SG, Sutton BP, Miquel ME, Nayak KS. Recommendations for real-time speech MRI. Journal of Magnetic Resonance Imaging 2016;43:28–44. https://doi.org/10.1002/JMRI.24997.
[36] Namasivayam AK, van Lieshout P. Speech motor skill and stuttering. J Mot Behav 2011;43:477–89. https://doi.org/10.1080/00222895.2011.628347.
[37] Packman A, Onslow M, Richard F, van Doorn J. Syllabic stress and variability: A model of stuttering. Clin Linguist Phon 1996;10:235–63. https://doi.org/10.3109/02699209608985174.