Oral-Reading Disfluency in Adolescents: An Exploratory Comparison of Choral Speech, Singing, and Metronome-Paced Speech

Teryn Rim[a]
[a] Singapore American School
40 Woodlands Street 41, Singapore 738547

Corresponding author: Teryn Rim
Email: terynrim@gmail.com
Phone: +65 8512 8097

Abstract

This exploratory study compared oral-reading disfluency and passage-completion rates across baseline reading, choral speech, singing, and metronome-paced speech in adolescents. Eight participants aged 16–18 completed ten English passages under each condition in a fixed order, yielding 320 recordings. Stuttering status was not clinically assessed. One researcher coded repetitions, prolongations, and blocks, and participant-level outcomes were summarized descriptively. Mean coded disfluency rates were 2.65 events per 100 target syllables at baseline, 0.74 during choral speech, 0.84 during singing, and 0.34 during metronome-paced speech. Corresponding mean passage-completion rates were 3.50, 3.14, 1.95, and 1.54 target syllables per second. All eight participants had lower disfluency rates during metronome-paced speech than at baseline. Choral speech combined lower mean disfluency with a completion rate closer to baseline than those in the other structured conditions. Fixed condition order, repeated exposure, the small sample, and unassessed coding reliability limit interpretation. These findings describe task-specific associations and do not establish treatment effects, lasting improvement, or transfer to everyday communication. Follow-up studies should counterbalance condition order, include rate-matched comparisons, and independently assess coding reliability.

Keywords: speech disfluency; choral speech; singing; metronome-paced speech; adolescents; speech rate; oral reading; rhythmic cueing; repeated measures; exploratory study.

1. Introduction

Stuttering is a speech disorder in which fluency is disrupted by sound and syllable repetitions, prolongations, and involuntary blocks that interrupt the forward movement of speech (1). These disruptions are behaviorally observable and measurable, making them trackable targets for research. Mawson et al. proposed a neuro-motor hypothesis of stuttering (2). The present study takes this behavioral definition as its foundation, focusing on quantifiable disfluency features rather than questions of etiology.

Stuttering in adolescence warrants attention because academic and social activities can place substantial demands on communication. Adolescence marks the transition from childhood to adulthood during which major physical transformations coincide with shifts in body image and the developing sense of self, while environmental pressures from peers, parents, teachers, and

society simultaneously increase (3). During this period of significant change, verbal demands intensify: oral presentations as part of the standard academic curriculum and social pressures in and outside of school further accumulate as part of the demands placed on communicative performance. Communication apprehension, social avoidance, and negative beliefs about communication have been reported among adolescents who stutter (3). These associations do not, by themselves, establish how particular academic or social pressures contribute to these experiences. The long-term risks associated with stuttering, including reduced social, psychological, and vocational opportunities and heightened vulnerability to social anxiety, may persist well into adulthood (3). These concerns provide a rationale for studying adolescent speech and communication, while the broader consequences of any task-related change in disfluency require separate investigation.

Music-based speech conditions provide one approach for examining task-related variation in disfluency. Wan et al. reviewed the neurological and behavioral relationships between singing and speech and discussed potential therapeutic applications of singing (4). The NIDCD also notes that singing or speaking in unison may temporarily reduce stuttering (1). These observations motivate comparisons of structured speech conditions, but do not establish their therapeutic value for adolescents or identify the mechanisms underlying any differences observed in the present study.

Broader music-therapy research has explored response mechanisms (5), cognition in Alzheimer's disease (6), depression in rat models (7), and attunement in young children with autism (8). These studies provide context, but their populations and outcomes do not directly establish effects on adolescent oral-reading disfluency.

The evidence base for non-pharmacological stuttering interventions also has limitations (9). Baxter et al. (2016), in a comprehensive systematic review, found evidence of positive outcomes across all types of interventions, with virtually all studies reporting at least some benefit for some participants (10). However, the authors also identified considerable individual variation in outcome, and the heterogeneous nature of the studies prevented meta-analysis. O'Donoghue et al. (2020) described treatment for adolescents who stutter as an under-researched area. This broader gap motivates exploratory work, but does not establish an absence of controlled research on music-based approaches (3).

The present study asks how coded disfluency rates and passage-completion rates differ across baseline reading, choral speech, singing, and metronome-paced speech in eight adolescent speakers. It aims to describe within-participant patterns across these conditions and identify procedural and measurement issues for follow-up research. Stuttering status was not clinically assessed, and the fixed condition order limits causal interpretation. The investigation therefore provides descriptive observations rather than evidence of a treatment effect or validation of a clinical research method.

2. Materials and Methods

2.1. *Study Design and Participants*

This exploratory study used a within-participant, repeated-measures design to compare oral-reading disfluency across four conditions: baseline reading, choral speech, singing, and metronome-paced speech. Eight participants aged 16–18 contributed data. The sample included native English speakers and bilingual speakers who used English as a second language. Stuttering status was not clinically assessed.

Consent was obtained electronically from participants and their parents or guardians through Google Forms. Participants were informed that study data would be retained for 12 months. Audio recordings were stored in Google Drive and were not shared with others. Participants were identified as P01–P08 in the reported results.

2.2. *Reading Materials*

Ten short English narrative passages were included in the analysis. The passages were prepared with ChatGPT assistance using an Aesop-style fable format and simplified vocabulary. Each participant read the same passage under all four conditions, keeping its wording and linguistic content constant within each comparison.

Target syllable counts were checked manually against the written passages. No pronunciation variants affecting these counts were identified. The ten passages contained 773 target syllables in total, averaging 77.3 syllables per passage. Additional syllables produced through repetition were not added to the target syllable count.

2.3. *Recording Procedure*

Recordings were collected between June and August 2026 using participants' personal recording devices, including phones and computers. No separate practice opportunities were provided. Participants recorded a single attempt for each passage-condition combination.

For each passage, participants completed the conditions in a fixed sequence: baseline reading, choral speech, singing, and metronome-paced speech. During baseline reading, participants read the passage aloud normally. During choral speech, they read simultaneously with a prerecorded spoken version of the passage. During singing, they sang along with a reference recording. The singing recordings used melodies adapted from classical music, and participants could sing in an octave comfortable for their voices. During metronome-paced speech, participants followed a tempo of 92 beats per minute, producing one target syllable per beat.

Each participant contributed ten recordings per condition, yielding 40 recordings per participant and 320 recordings overall. Condition order was neither randomized nor counterbalanced.

2.4. *Disfluency Coding*

The author manually coded all recordings while following the corresponding passage text. Disfluencies were classified as repetitions, prolongations, or blocks. Repetitions involved repeated speech sounds or syllables; prolongations involved an abnormal extension of a speech sound; and blocks involved an apparent interruption in attempted speech production.

Each identified disfluency event was counted once under its corresponding category. Fillers, revisions, and ordinary pauses were excluded. Intended prolongations required by the singing melody were not coded as disfluent prolongations. No second coder independently evaluated the recordings, and inter-rater reliability was not assessed.

2.5. *Outcome Measures*

The primary outcome was the coded disfluency rate, expressed as events per 100 target syllables. For each recording, repetition, prolongation, and block counts were summed, divided by the passage's target syllable count, and multiplied by 100.

Passage-completion rate was calculated by dividing the target syllable count by reading duration in seconds. Initial and final silence were excluded from the duration measurement, while pauses occurring within the passage were included. This measure represented overall passage-completion pace, including internal pauses.

2.6. *Data Analysis*

The analysis was descriptive. For each participant and condition, disfluency counts were summed across the ten passages and divided by the corresponding total of 773 target syllables, then multiplied by 100. Participant-level passage-completion rates were calculated by dividing total target syllables by the summed duration of the ten recordings.

Condition-level means and sample standard deviations were calculated across the eight participant-level values, giving each participant equal weight in the group summaries. Repetitions, prolongations, and blocks were also totaled separately by condition.

Absolute reductions from baseline were calculated by subtracting each condition mean from the baseline mean. Relative reductions were calculated by dividing this difference by the baseline mean and multiplying by 100. The 320 recordings were treated as repeated observations within eight participants rather than as independent participant observations.

3. Results

The dataset comprised 320 reading records from eight participants, with ten passages per participant in each of four conditions: baseline reading, choral speech, singing, and

metronome-paced speech. All records were included, yielding 80 records and 6,184 target syllables per condition. Results were summarized descriptively across the eight participants.

The mean coded disfluency rate was 2.65 events per 100 target syllables at baseline, compared with 0.74 during choral speech, 0.84 during singing, and 0.34 during metronome-paced speech (Table 1). Metronome-paced speech had the lowest observed mean. The mean rates for choral speech, singing, and metronome-paced speech were lower than baseline by 1.91, 1.81, and 2.31 events per 100 target syllables, respectively. Relative to the baseline mean, these rates were 72.0%, 68.3%, and 87.2% lower.

Mean syllable rate was 3.50 target syllables per second at baseline, 3.14 during choral speech, 1.95 during singing, and 1.54 during metronome-paced speech (Table 1). The mean rate was 10.3% lower than baseline during choral speech, compared with 44.4% during singing and 55.9% during metronome-paced speech. The condition with the lowest mean coded disfluency rate also had the lowest mean syllable rate.

All eight participants had a lower coded disfluency rate during metronome-paced speech than at baseline. Individual responses to singing were less consistent: six participants had lower rates, one had an unchanged rate, and one had a higher rate. P02 had five baseline events and six singing events across 773 target syllables in each condition, corresponding to 0.65 and 0.78 events per 100 syllables. P07 had five events in both conditions, equivalent to 0.65 events per 100 syllables in each.

The recorded event totals were 164 at baseline, 46 during choral speech, 52 during singing, and 21 during metronome-paced speech (Table 2). Blocks were the largest category in each condition, with 77, 29, 29, and 9 events, respectively. The entered counts of repetitions, prolongations, and blocks were each lower in all three structured conditions than at baseline.

Table 1. Coded disfluency and passage-completion rates by condition

Condition	Records	Coded events	Disfluency rate, mean (SD)	Completion rate, mean (SD)
Baseline	80	164	2.65 (1.98)	3.50 (0.48)
Choral speech	80	46	0.74 (0.57)	3.14 (0.09)
Singing	80	52	0.84 (0.25)	1.95 (0.05)
Metronome-paced speech	80	21	0.34 (0.29)	1.54 (0.02)

Note. Disfluency rates are coded events per 100 target syllables; completion rates are target syllables per second. Values are means (SD) across eight participants. Each condition includes 80 recordings and 6,184 target syllables.

Table 2. Recorded counts of coded disfluency types by condition

Condition	Repetitions	Prolongations	Blocks	Total
Baseline	50	37	77	164

Condition	Repetitions	Prolongations	Blocks	Total
Choral speech	12	5	29	46
Singing	12	11	29	52
Metronome-paced speech	5	7	9	21

Note. Counts are pooled across eight participants and ten passages per condition; each condition includes 80 recordings and 6,184 target syllables.

4. Discussion

All three structured conditions were associated with a lower mean disfluency rate than baseline reading. The reduction was observed across every disfluency type recorded: repetitions, prolongations, and blocks each declined in choral speech, singing, and metronome-paced speech compared with the baseline reading. Metronome-paced speech showed the largest reduction in the group mean, and every participant had a lower disfluency rate than at baseline. Choral speech showed a reduction in mean disfluency rate similar to that observed during singing, alongside a syllable rate much closer to baseline than in the other two structured conditions.

Changes in auditory input, rhythm, timing, and coordination offer possible explanations for the lower disfluency rate. The conditions combined features that could be examined as possible explanations in future research. Choral speech provided an accompanying recorded voice; singing combined an accompanying voice with melody and rhythm; and metronome-paced speech provided regular timing cues. The present study did not isolate these features or test their mechanisms. Differences in passage-completion rate and the fixed order of repeated readings also limit interpretation.

Meekings et al. found different rhythmic responses to choral speech in adults who stutter and neurotypical speakers. These findings provide limited support for the specific hypothesis that choral speech compensates for a motor timing deficit through rhythmic alignment (11).

Calculated from the group means, the relative reduction in coded disfluency rate was greatest during metronome-paced speech (87.2%), followed by choral speech (72.0%) and singing (68.3%). The corresponding reductions in syllable rate were 55.9%, 10.3%, and 44.4%, respectively. These percentages describe differences relative to baseline; they do not establish statistically significant differences between conditions.

Choral speech was notable because its lower coded disfluency rate accompanied a smaller change in syllable rate than occurred during singing or metronome-paced speech. This combination makes choral speech a useful focus for further investigation. Speech naturalness, effort, intelligibility, and participant preference were not measured, and baseline reading itself was a structured task.

The singing condition combined melody, sustained voicing, an accompanying recorded voice, and a slower passage-completion rate. Their contributions cannot be separated from one another or from repeated exposure in this design. Metronome-paced speech also imposed a strong

constraint on timing. The instruction to produce one syllable per beat at 92 beats per minute specified a target of 1.53 syllables per second, close to the observed mean of 1.54. Comparisons across conditions remain descriptively useful, but they cannot isolate the contribution of melody or external timing from that of rate. Previous experimental work on metronome-paced speech likewise demonstrates the importance of controlling rate when investigating speech-production mechanisms (12).

For a fixed passage length, lower passage-completion rates imply longer reading times. The markedly lower completion rates in singing and metronome-paced speech should therefore be considered alongside their lower disfluency rates. Lower coded disfluency alone does not establish more natural, less effortful, or more effective communication, because these outcomes were not assessed. Choral speech was accompanied by a smaller rate difference from baseline, but this does not demonstrate an advantage in everyday communication.

Individual responses to singing varied. P02 had five baseline events and six singing events across 773 syllables, and P07 had five events in both conditions. A participant can have a higher value than other participants without having increased relative to their own baseline. P02's singing rate of 0.78 was below the group singing mean of 0.84, yet it was above P02's own baseline rate of 0.65. These differences also rest on very small counts. P02's change amounted to a single event; the present design and coding procedure cannot distinguish a condition-related difference from ordinary variation or coding uncertainty. Every participant had a lower rate during metronome-paced speech than at baseline, showing a consistent direction of difference across this sample.

The pooled condition summaries do not show whether performance changed across weeks. A week-level analysis would need to consider passage differences, familiarity with the procedure, and ordinary variability in disfluency. The results as currently summarized do not establish sustained improvement.

Blocks showed the largest absolute reduction in every condition, falling by 48 events in choral speech and singing and by 68 in metronome-paced speech. Proportionally, the largest reductions differed by condition: prolongations in choral speech (86.5%), repetitions in singing (76.0%), and repetitions in metronome-paced speech (90.0%). Blocks were also the most frequent baseline category (77 events), so they had the most room to decrease in absolute terms. The proportional reduction in blocks was smaller than that in the other categories in choral speech (62.3%) and singing (62.3%). This describes the observed counts rather than establishing that blocks are more resistant to these conditions. Because the baseline frequencies differ, absolute and proportional reductions answer different questions and should be read together.

Repeated exposure is a plausible alternative explanation. Using the same passage in all four conditions held the text constant, so differences in vocabulary, sentence structure, and word difficulty could not account for differences between conditions within a passage. However, familiarity could still change across readings. By the second, third, and fourth reading, participants may have known the pronunciation of difficult words, anticipated upcoming sentence structure, and read more smoothly for reasons unrelated to the condition. Because the

order was fixed, with baseline first and then choral speech, singing, and metronome-paced speech, any effect of repeated exposure is confounded with the effect of condition. It is impossible to separate the two with this design. The metronome condition was always read last, so its lower disfluency rate is also compatible with an order-related explanation. This study did not test whether familiarity increased or accounted for the difference.

Classifying repetitions, prolongations, and blocks requires judgement. Although ordinary pauses were excluded, distinguishing a block from a pause can be uncertain. Intended prolongations required by the singing melody were also excluded, but distinguishing them from disfluent prolongations still required judgement. An accompanying recorded voice in choral speech or singing may obscure the participant's own speech. All recordings were coded by one researcher, and inter-rater reliability was not assessed. The consistency of event identification and classification is therefore uncertain, and condition-specific recording features may have affected the counts.

Each passage contained about 77 syllables on average, so one event corresponds to roughly 1.3 events per 100 syllables in an individual passage. The reported participant-level rates, however, pooled 773 target syllables across ten passages; one event changes that pooled rate by approximately 0.13 events per 100 target syllables. This distinction is important when interpreting small count differences such as P02's five baseline events and six singing events. Low baseline counts also leave limited room for a further decrease. Numerical differences should be distinguished from practical importance: naturalness, listener perceptions, and the speaker's experience were not assessed, so the communication value of the observed reductions cannot be determined.

The sample of eight participants provides a limited basis for generalization, and the repeated recordings are not independent observations from different speakers. Stuttering status was not clinically assessed, so the sample cannot be characterized as a clinically diagnosed stuttering population. Prepared reading also differs from everyday communication, which can involve unplanned wording, listener interaction, and emotional pressure. The results describe immediate performance on the recorded tasks. They do not establish a lasting treatment effect or transfer to everyday communication.

Choral speech warrants further investigation because its lower mean coded disfluency rate accompanied a passage-completion rate closer to baseline than those in singing or metronome-paced speech. This pattern does not establish that choral speech itself caused the difference. The study produced recordings and descriptive summaries, but did not independently establish coding reliability or clinical usefulness. A follow-up study should randomize or counterbalance condition order to address repeated exposure, and include rate-matched comparisons to examine whether condition-related differences remain at comparable completion rates. Independent coding with an agreement assessment would help evaluate measurement reliability. Measures of naturalness, effort, intelligibility, and participant preference would help determine whether any observed differences have practical importance.

5. Conclusion

Across eight adolescent speakers, choral speech, singing, and metronome-paced speech were each associated with lower mean coded disfluency than baseline reading. Metronome-paced speech had the lowest mean disfluency rate and the slowest passage-completion rate. Choral speech combined a lower disfluency rate with a completion rate closer to baseline than those in the other structured conditions. These observations support further investigation of choral speech under more controlled conditions. Fixed condition order and differences in completion rate limit causal interpretation, while unassessed coding reliability adds measurement uncertainty. Stuttering status was not clinically assessed, and the findings do not establish treatment effects, lasting improvement, or transfer to everyday communication. Follow-up studies should control condition order and rate and independently assess coding reliability.

AI use disclosure: ChatGPT (OpenAI) was used to assist with manuscript drafting and revision, reference formatting, and verification of descriptive calculations. The author is responsible for the accuracy and integrity of the final manuscript.

6. References

1. National Institute on Deafness and Other Communication Disorders. (2017, March 6). *Stuttering.* https://www.nidcd.nih.gov/health/stuttering
2. Mawson, A. R., Radford, N. T., & Jacob, B. (2016). Toward a theory of stuttering. *European Neurology, 76*(5–6), 244–251. https://doi.org/10.1159/000452215
3. O'Donoghue, J., Moss, H., Clements-Cortes, A., & Freeley, C. (2020). Therapist and individual experiences and perceptions of music therapy for adolescents who stutter: A qualitative exploration. *Nordic Journal of Music Therapy, 29*(4), 353–370. https://doi.org/10.1080/08098131.2020.1745872
4. Wan, C. Y., Rüber, T., Hohmann, A., & Schlaug, G. (2010). The therapeutic effects of singing in neurological disorders. *Music Perception, 27*(4), 287–295. https://doi.org/10.1525/mp.2010.27.4.287
5. Clements-Cortes, A., & Bartel, L. (2018). Are we doing more than we know? Possible mechanisms of response to music therapy. *Frontiers in Medicine, 5*, Article 255. https://doi.org/10.3389/fmed.2018.00255
6. Fang, R., Ye, S., Huangfu, J., & Calimag, D. P. (2017). Music therapy is a potential intervention for cognition of Alzheimer's disease: A mini-review. *Translational Neurodegeneration, 6*, Article 2. https://doi.org/10.1186/s40035-017-0073-9
7. Le, J., Deng, W., & Le, T. (2025). Music therapy in depression: Exploring mechanisms and efficacy in rat models. *Brain Sciences, 15*(4), Article 338. https://doi.org/10.3390/brainsci15040338
8. Mössler, K., Schmid, W., Aßmus, J., Fusar-Poli, L., & Gold, C. (2020). Attunement in music therapy for young children with autism: Revisiting qualities of relationship as mechanisms of change. *Journal of Autism and Developmental Disorders, 50*(11), 3921–3934. https://doi.org/10.1007/s10803-020-04448-w

9. Baxter, S., Johnson, M., Blank, L., Cantrell, A., Brumfitt, S., Enderby, P., & Goyder, E. (2015). The state of the art in non-pharmacological interventions for developmental stuttering. Part 1: A systematic review of effectiveness. *International Journal of Language & Communication Disorders, 50*(5), 676–718. https://doi.org/10.1111/1460-6984.12171
10. Baxter, S., Johnson, M., Blank, L., Cantrell, A., Brumfitt, S., Enderby, P., & Goyder, E. (2016). Non-pharmacological treatments for stuttering in children and adults: A systematic review and evaluation of clinical effectiveness, and exploration of barriers to successful outcomes. *Health Technology Assessment, 20*(2), 1–302. https://doi.org/10.3310/hta20020
11. Meekings, S., Eijk, L., Stankova, S., Maruthy, S., & Scott, S. K. (2025). Conflicting evidence for a motor timing theory of stuttering: Choral speech changes the rhythm of both neurotypical and stuttering talkers, but in opposite directions. *Journal of Speech, Language, and Hearing Research, 68*(7S), 3558–3567. https://doi.org/10.1044/2024_JSLHR-24-00405
12. Davidow, J. H. (2014). Systematic studies of modified vocalization: The effect of speech rate on speech production measures during metronome-paced speech in persons who stutter. *International Journal of Language & Communication Disorders, 49*(1), 100–112. https://doi.org/10.1111/1460-6984.12050