The Relationship between Strum Rhythm and Speech Pauses in a Trial Study of Music Enriched Verb Network Strengthening Treatment

Kayla M. Schoolar

Music Therapy Department

First Reader: Dr. Gayle DeDe, Speech-Language Department

Second Reader: Dr. Lillian Eyre, Music Therapy Department

May 5, 2022

Abstract

Aphasia is a language disorder that may be caused by stroke or traumatic brain injury. Individuals with aphasia (IWAs) may experience difficulties with word finding, reading, auditory comprehension, writing, and/or other speech and language related challenges. Many IWAs receive speech-language therapy to address these issues. One speech-language intervention for this population is Verb Network Strengthening Treatment (VNeST), which focuses on action verbs and related subject-direct object (agent-patient) pairs. Music therapy (MT) has also been used to treat aphasia, exploiting elements common to speech and music such as pitch, rhythm, and prosody to address non-musical clinical goals related to verbal expression.

MeVNeST (Music Enriched Verb Network Strengthening Treatment) is an experimental music therapy treatment currently being investigated by doctoral candidate Jing-Wen Zhang. MeVNeST follows the same steps as traditional VNEST but replaces spoken prompts and responses with improvised singing in the context of live, improvised guitar accompaniment. The purpose of both VNEST and MeVNEST is to improve clients' ability to retrieve common action verbs along with related nouns.

The present study focuses on the relationship between accented beats in the guitar accompaniment of MeVNeST sessions and the duration of pauses before and within client responses to clinician prompts. The goal is to determine whether certain strum patterns are associated with more efficient word retrieval. Two strum conditions were identified: a normative beat (unsyncopated) and a syncopated beat. Fourteen audio samples from Zhang's study were considered: seven from each strum condition, drawn from two study participants. Music and speech data from two of these samples were then coded in Praat and analyzed in Excel.

Results showed no statistically significant difference between strum conditions for pauses before responses. There was a statistically significant pause duration within responses, with shorter responses being associated with the syncopated strum condition. Future studies should continue to examine how rhythmic musical accompaniment might be used to support and enhance purposeful, fluent speech and to minimize verbal hesitations.

Keywords: Music therapy, aphasia, rhythm, neurologic music therapy, speech language therapy

Introduction

According to the National Aphasia Association, aphasia is “an acquired communication disorder that impairs a person’s ability to process language but does not affect intelligence. Aphasia impairs the ability to speak and understand others” (National Aphasia Association, n.d.). Over two million people in America experience aphasia (National Aphasia Association, n.d.). Aphasia is often caused by strokes, and as medical technology improves and stroke survival rates increase, the number of individuals with aphasia (IWA) is expected to increase (Lim et al., 2013). IWAs may experience difficulties with word finding, sentence formation, and comprehension of written and spoken language as a result of acquired brain damage (i.e., a stroke or traumatic brain injury). IWAs may also experience *apraxia of speech*, which impacts oral motor planning, or *dysarthria*, a weakening or loss of coordination of the muscles used in speech. Individuals with apraxia may be inconsistent in their word pronunciation and/or articulation, while individuals with dysarthria tend to have quiet, breathy voices (Hobson, 2006).

Literature Review

Treatments

IWAs may participate in speech therapy to treat aphasia symptoms. The National Aphasia Association mentions several common therapeutic techniques (Davis, 2011). Constraint-induced therapy requires the IWA to practice using an impaired speech or language function without compensatory strategies. For instance, a person who struggles to retrieve (remember and say) verbs may rely on pantomime to communicate their point. Constraint-induced therapy would encourage them to describe the action verbally instead. IWAs may also be given pictures to describe, scripted conversations to practice, and unscripted conversation practice either in support groups with other IWAs or with trained non-aphasic volunteers or therapists.

Music Therapy Treatments

Some IWAs may also elect to participate in Music Therapy (MT). One particularly common MT method with IWAs is Neurologic Music Therapy (NMT), “a method of music therapy for improving the cognitive, sensory, and motor functions impaired by neurological diseases” (Lim et al., 2013, p. 557). According to Thaut (2014), NMT seeks to use music’s ability to stimulate multiple areas and processes in the brain, promoting neuroplasticity. Lee et al. (2018) lists eight standardized NMT techniques for treating aphasia: a) Melodic Intonation Therapy (MIT), Musical Speech Stimulation (MUSTIM), Rhythmic Speech Cuing (RSC), Vocal Intonation Therapy (VIT), Oral Motor and Respiratory Exercises (OMREX), Therapeutic Singing (TS), Developmental Speech and Language Training Through Music (DSLM), and Symbolic Communication Training through Music (SYCOM).

Of these eight, MIT has received considerable attention in aphasia-related literature. MIT involves singing a common phrase or sentence on two notes with a slow rhythm. Accented syllables are sung on a higher pitch and are held for a full beat while unaccented syllables are sung on a lower pitch and held for half a beat. Various strategies such as tapping one hand in rhythm or speaking in unison with the clinician may be initially incorporated and gradually withdrawn until the client is able to sing the phrase or sentence independently (Zumbansen et al., 2014). In a study of three native French-speaking men, all of whom had aphasia, participation in MIT was associated with greater generalizability of trained sentences than were the control condition or engagement in rhythmic speech without melody. MIT may be implemented by both speech-and-language pathologists and music therapists, making it a versatile treatment technique.

Singing is also a common technique used by music therapists with IWAs. Lim et al. (2013) notes that singing may provide IWAs with more time to produce syllables and phonemes.

Singing may also assist IWAs who experience dysarthria along with aphasia, since it requires breath support, the ability to produce a variety of pitches, and may be used to increase vocal intensity (Hobson, 2006).

Importance of Rhythmic Elements

Although the melodic component of music is important to the use of music therapy with IWAs, a growing body of research is focusing on rhythm as an impetus for speech improvement (Lee 2018). A study by Zumbansen et al. (2014) argues that the combination of pitch and rhythm is responsible for MIT's efficacy, since participants showed greater improvement following MIT than following unpitched rhythmic speech interventions. Stahl et al. (2011) asserts, however, that of these two factors, rhythm is the more important. In a study of seventeen IWAs in Berlin, Germany, pitch by itself was not found to have an effect on speech production, whereas rhythmic speech supported the correct production of spoken syllables, especially among participants with damage to the basal ganglia. According to Stahl (2011), then, singing is primarily helpful because it is "rhythm in disguise" (title)—an enjoyable way of rhythmically structuring speech.

The Sound Envelope Processing (SEP) hypothesis posited by Fuji and Wan (2014) also points to rhythm as a key element of effective speech, arguing that entrainment to a pulse may stimulate the neural networks central to speech and communication. Based on this stance, they suggest aphasia as one of the diagnoses that might benefit from a rhythm-based therapy. This focus on rhythmic entrainment is part of the theoretical framework of NMT (Thaut 2015). Entrainment is the phenomenon in which two systems' motion or frequency synchronize with each other; a musical example of this would be the tendency of persons listening to or playing music together to move, play, or sing in unison.

Music Therapy Foundational Study

Verb Network Strengthening Treatment (VNeST) is a multi-step intervention used by speech-language pathologists to promote IWAs' ability to retrieve common action verbs as well as associated subjects (agents) and direct objects (patients). During a VNeST session, clients are shown three cards labelled *who, what,* and **verb in question* *. The clinician prompts the client by asking "Who might throw something?" or "Who sews what?" and helps the client to generate several agent-patient pairs. The pairs are written out as the client says them. Once several pairs have been chosen and written down, the client reads them aloud—with help from the clinician if needed—and chooses one scenario to expand by asking *where, when,* and *why* questions. For instance, "The grandmother sews the shirt *in the living room/ after dinner/ because it has a hole in it.*" Again, responses are written down and the client is asked to read the full sentence. All written materials are then removed, and the client is asked to listen to a series of sentences using the verb in question and to decide if the verb is used correctly or incorrectly (i.e., agent or patient is inappropriate or agent and patient are reversed). After completing this task, the clinician asks the client which verb they have been working on. Finally, the client repeats step one (generating agent-patient pairs) without cues until they are unable to think of any more pairs. Depending on the length of the session, this process may be repeated for several verbs (Edmonds, 2018).

Music Enriched Verb Network Strengthening Treatment (MeVNeST) is an experimental treatment based on VNeST (Zhang, dissertation in progress). MeVNeST follows the same steps as a traditional VNeST session, but with an additional musical element. Rather than relying on spoken dialogue between a speech-language pathologist and the participant, the co-facilitating music therapist provides an improvised accompaniment on guitar and asks the client to generate a short, wordless melody. Prompts and responses are then sung using this melody. The music therapist uses common harmonic progressions to cue responses. For instance, when asking

“Which verb are we working on?” the music therapist plays a dominant chord, the musical equivalent of an expectant expression, and then resolves to a tonic chord, the musical equivalent of a closure or resolution, as the client names the verb.

Purpose

This study seeks to examine the relationship between the rhythm of the guitar accompaniment during MeVNeST sessions and the duration of pauses before and within client responses. Pauses may indicate that an IWA is struggling to generate a response. Pauses may occur before a response (i.e. an IWA hesitates before beginning to speak) or within a response (i.e. an IWA pauses between words in a phrase or sentence or pauses within a word).

Rhythm can be used to structure and support speech production during music therapy with IWAs. Existing MIT protocols demonstrate precedent for the use of a steady metronomic beat in conjunction with rhythmic speech. Phenomenologically, the predictability of steady quarter notes may provide an organizational framework for speech production while demanding less of an IWA’s attention than a syncopated rhythm would.

The long-term goal of this work is to understand the practical use of rhythm in music therapy with IWAs and to determine whether different rhythms are associated with greater treatment effects. Knowledge of which rhythms, if any, are most conducive to smooth speech production would allow music therapists who work with these clients to provide better-informed, more effective treatment tailored to their clients’ unique speech and language needs.

In order to examine the relationship between accompaniment rhythm and participant speech fluency (as measured by pause duration), this study posited the following questions:

1. Does pause duration before participant responses differ with strum condition (normative vs. syncopated)?

2. Does pause duration within participant responses differ with strum condition (normative vs. syncopated)?

3. Does pause duration before participant responses change over the course of treatment?

4. Does pause duration within participant responses change over the course of treatment?

The present study was designed to permit examination of treatment effects over time. However, data coding and analysis focused on one sample from each strum condition to evaluate the feasibility of the coding method. Thus, the present paper focuses on questions 1 and 2.

The hypothesis of this study is that a normative strum pattern characterized by greater predictability and less syncopation will be positively related to shorter pause durations, both before and within responses.

Methods

Participants

IRB approval for this study was granted by Temple IRB, and all participants provided informed consent. Two participants were selected from Zhang's study (dissertation in progress). Zhang studied four White males over age fifty with moderate to severe aphasia following a stroke. All participants were in the chronic stage of aphasia (over one year since onset) and presented with relatively intact auditory comprehension. Participants also exhibited agrammaticism (i.e., omitting verbs and important function words such as prepositions). Two participants (P1 and P2) participated in a combination of VNeST and MeVNeST sessions (seven and seven), and the other two (P3 and P4) participated only in MeVNeST sessions (fourteen total). Data for this study were drawn from P3 and P4's sessions.

In order to compare the fluidity of speech (as measured by the duration of pauses within participant responses) at the beginning and end of the MeVNeST treatment period, a substantial amount of time was needed between the initial and final MeVNeST sessions. A sizable number of sessions was also needed to identify multiple instances of each strum condition per participant. For these reasons, samples were drawn only from sessions with P3 and P4, who participated in fourteen MeVNeST sessions each. P1 and P2 participated in only seven MeVNeST sessions, which was not enough time to influence speech fluidity and did not provide a large enough pool from which to draw samples of the critical strum patterns.

Sample Selection

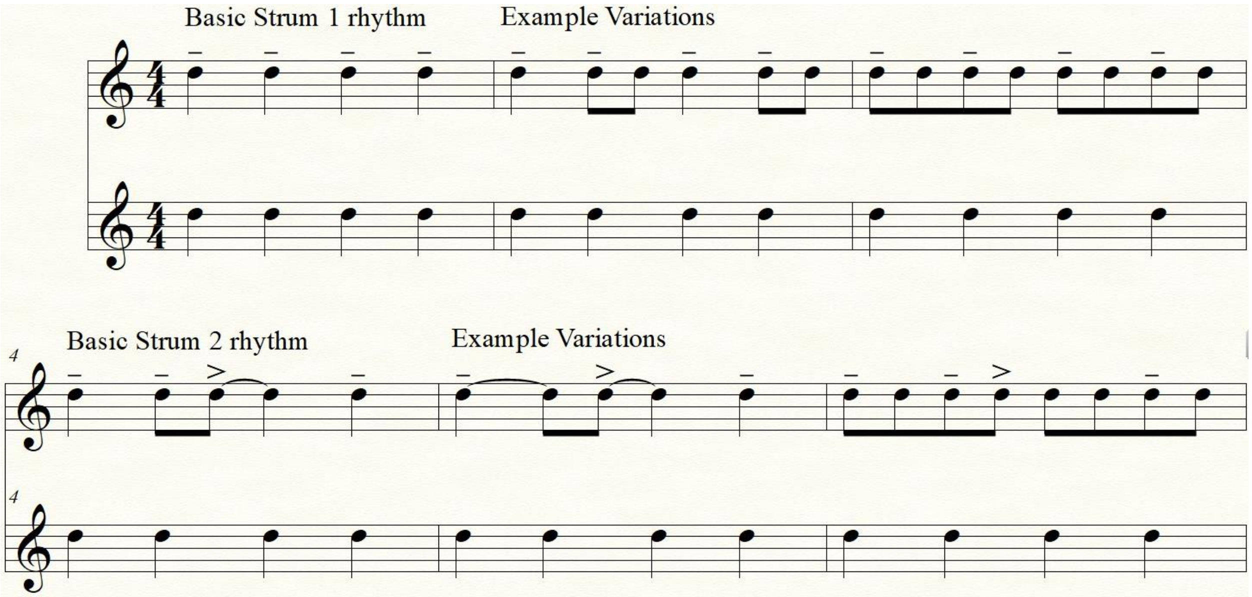
Because Step 1 of the MeVNeST procedure (generating agent-patient pairs) required the most independently generated speech from the participants, samples for analysis were selected from this step. Before defining strum conditions or choosing samples, the author reviewed the video recordings of fourteen sessions with P3 and P4's (see Table 1 on p. 12) and transcribed the guitar strum pattern during Step 1 of each verb. Participants typically trained 3-5 verbs per MeVNeST session, so a total of 53 samples were transcribed from the fourteen sessions.

Based on these transcriptions, two strum conditions were identified: a normative condition (Strum 1) and a syncopated condition (Strum 2). Strum 1 was defined as a strum in four-four time (four beats per measure) with a downstroke on every quarter note beat, regardless of where the accented beats fell. The strum needed to be steady and predictable, with no displaced beats. Strum 2 was defined as a strum in four-four time that had a displaced third beat, preferably with no other syncopations. The strum needed to be consistent while maintaining a sense of syncopation on the displaced third beat.

These two strums were chosen for their respective predictability. Traditionally, steady quarter notes are considered the default in four-four time. Listeners instinctively expect to hear a stroke on each beat. Syncopation subverts this expectation, disrupting the predictability of the music and creating a jarring emphasis on the displaced beat (Fig. 1).

Fig. 1

Comparison of Strum Conditions and Variations



Note: Tenuto [-] markings show the four basic beats of each measure. Accent [>] markings show Strum 2's displaced third beat. A continuous quarter note rhythm (identical to the basic Strum 1 rhythm) has been provided for reference.

In choosing samples for analysis, sessions were divided into four blocks: Block 1 (sessions 1-4), Block 2 (session 5-7), Block 3 (sessions 8-10) and Block 4 (sessions 11-14). To compare pause duration over time, an equal number of samples were chosen from Blocks 2 and 4. No samples were chosen from Block 1 under the assumption that participants were adjusting to the treatment format. To maximize time between earlier and later samples, no sessions were chosen from Block 3. Two samples were chosen per client, strum condition, and session Block for a total of fourteen samples. Ideally, sixteen samples would have been chosen, but only one satisfactory example of Strum 1 was found for P4 in Block 2, and only one satisfactory example of Strum 2 was found for P3 in Block 4 (See Fig. 2).

Table 1

Block and Sample Divisions

	Block 1				**Block 2**			**Block 3**				**Block 4**		
Session	1	2	3	4	5	6	7	8	9	10	11	12	13	14
	Too early in treatment-no samples drawn from this block.				Early treatment samples			No samples drawn from this block to provide space between earlier and later samples.				Late treatment samples		
Participant 3	0 samples				Total of 4 samples Strum 1: 2 Strum 2: 2			0 samples				*Total of 3 samples* Strum 1: 2 *Strum 2: 1*		
Participant 4	0 samples				*Total of 3 samples* *Strum 1: 1* Strum 2: 2			0 samples				Total of 4 samples Strum 1: 2 Strum 2: 2		

A full analysis of all fourteen samples would be outside the scope this study due to time limitations. Instead, two samples from Block 2, one from each strum condition, were selected for close analysis. These samples were chosen because of how faithful the guitar rhythm in each was to the established strum conditions (Fig. 2, 3).

Fig. 2

Strum 1: Transcription Examples of Guitar and Dialogue

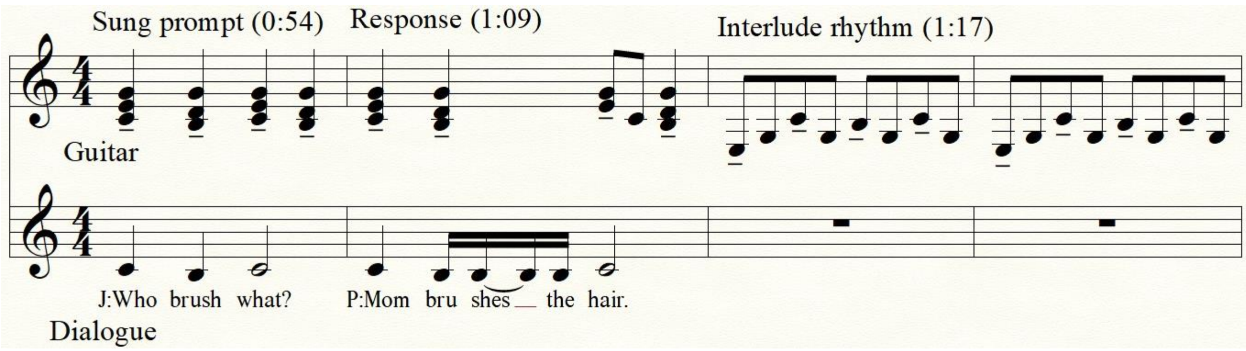


Fig. 3

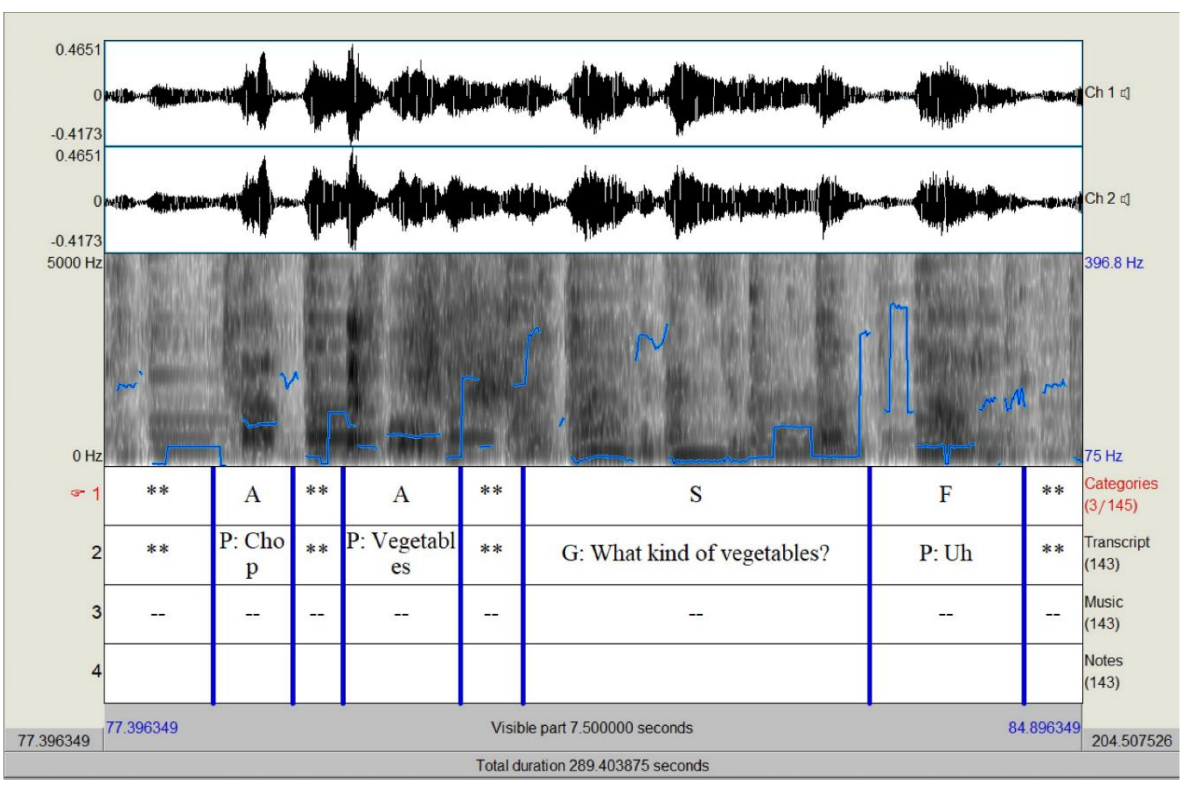
Strum 1: Transcription Examples of Guitar and Dialogue

Coding and Data Analysis

Samples were coded in the computer program Praat (Boersma and Weenink, 2021), which allows for annotation of audio files. Four annotation tiers were used: *Transcript*, *Music*, *Categories*, and *Notes*. The *Transcript* tier contained a direct audio transcription, including silences and filler sounds. The *Music* tier included a running description of the guitar rhythm, noting any inconsistencies and changes. Expected entrances were also noted here (i.e., the point at which the participant would be expected to begin singing). The *Categories* tier contained an analysis of the transcript data. Pauses, filler sounds, prompts, and responses were noted here. Any perceptible silence before a participant response or within a participant response was considered a pause. A full list of coding abbreviations and definitions in included in Appendix A. The *Notes* tier housed miscellaneous observations that clarified information in the other tiers.

Fig. 3

Example of Praat Coding System



Note: This figure shows an example of how the coding tiers were laid out in Praat. Not all coding terms appear here. For a full explanation of coding terms, see the Appendix.

Textgrid data was imported into Excel. Pauses were measured from the offset of a clinician prompt or participant response to the onset of the following participant response (pause before response) and from the offset of one word in a participant response to the onset of the next word (pause within response). Filler sounds were considered pauses. All other participant responses, whether MeVNeST responses or otherwise, were considered content words, and pauses were measured from their off/onset timestamps. In cases where there was no silence between participant words or between clinician and participant responses, a value of zero (0) was assigned. The mean and standard deviation of each type of pause (before or within response and for each strum condition) was calculated, as well as the standard deviation. Pause duration between and within responses were analyzed in separate t-tests.

Results

The normative Strum 1 sample lasted 371 seconds (6.2 minutes). In that time, the participant generated five pairs for an average of one pair every 74 seconds. The syncopated Strum 2 sample lasted 283 seconds (4.7 minutes), and the participant generated four pairs for an average of one pair per 71 seconds.

The mean pause length for the Strum 1 (normative) sample was 3.888 seconds before responses with a standard deviation of 4.065 seconds and 1.427 seconds within responses with a standard deviation of 2.73 seconds. The median pause length before responses was 2.435 seconds, and the median pause length within responses was 0.381 seconds. For the Strum 2 (syncopated) sample, the mean pause length before responses was 6.214 seconds with a standard deviation of 7.205 seconds and 0.848 seconds within responses with a standard deviation of 1.430 seconds. The median pause length before responses was 3.582 seconds, and the median pause length within responses was 0.320 seconds

Outliers were operationally defined as pause durations greater than two standard deviations from the mean for each strum condition and pause location. Outliers were removed to reduce the possibility of including atypically long pauses that reflected external interruptions (e.g., a particularly long pause due to a phone call or a written clinician cue that was not detectable in the audio sample). Five data points were removed from the Strum 2 sample—two from the "before" pauses and two from the "within" pauses. Five data points were also removed from the Strum 1 sample—two from the "before" pauses and two from the "within" pauses. In total, 93 data points were considered from the Strum 1 sample, 34 "before" and 59 "within", and 48 data points were considered from the Strum 2 sample, 18 "before" and 30 "within".

Means and standard deviations were then calculated based on the cleaned data (see Table 2, 3). The Strum 1 condition showed a 3.261 second mean for pauses before responses with a standard deviation of 3.188 seconds. There was a 0.935 second mean for pauses within responses with a standard deviation of 1.560 seconds. The Strum 2 condition showed a mean pause duration of 4.178 seconds before responses with a standard deviation of 3.225 seconds. The mean pause duration within responses was 0.450 seconds with a standard deviation of 0.075 seconds. T-tests showed no significant effect of strum type on pauses before responses, $t(31)=0.89$, $p=0.37$. However, the effect of strum type was significant for pauses within responses, $t(84)=2.07$, $p=0.04$.

Table 2 ***Pauses Before Responses***

	Strum 1	*Strum 2*
Mean	3.261	4.178
Standard deviation	3.188	3.225
Number of data points	34	18

t(31)=-0.89

p=0.37

Table 3

Pauses Within Responses

	Strum 1	*Strum 2*
Mean	0.935	0.450
Standard deviation	1.560	0.075
Number of data points	59	30

t(84)=2.07

p=0.04

Discussion

These results should be interpreted with caution given that data from only one individual are included in the analysis and that there were unequal numbers of observations across conditions. Additionally, data was drawn from only two of 14 possible samples for this participant. Nonetheless, these results provide preliminary evidence that the Strum 2 condition (syncopation on third beat of the measure) is associated with shorter pause durations within participant responses than the Strum 2 condition.

The results of this study show a statistically significant decrease in pause duration within responses under Strum 2 conditions. This is contrary to the initial hypothesis that the un-syncopated rhythm of Strum 1 would be more conducive to uninterrupted speech production. No statistically significant difference between strum conditions was found in pause duration before responses. This is also contrary to the initial hypothesis that the predictable rhythm of Strum 1 would be associated with shorter pauses before responses.

There are several possible explanations for these findings. Rhythmically, the syncopation of the Strum 2 condition may have helped the participant to orient himself within the music by audibly differentiating between beats. Within the Strum 2 sample, the participant tended to speak in a syncopated rhythm that matched that of the strum condition. This may indicate that he found the syncopated rhythm easier to replicate than straight quarter notes. If this is the case, it should be noted that this may be an individual preference and may not hold true for all IWAs.

Rhythmically based speech therapy interventions have tended to rely on straight, metronomic beats to support client speech production. In MIT, for example, clients practice saying common phrases while tapping an even, unmetered (not divided into measures) foundational beat. Natural speech, however, uses a variety of rhythms. Some syllables and words are held longer than others; words may be flanked by silence or may connect to the next word without a pause. MIT protocol reflects this organic rhythm in that clients are taught to hold accented syllables twice as long as unaccented syllables. It is possible, however, that using a base rhythm that reflects the natural rhythm of speech—a syncopated rhythm such as the one in Strum 2, for example, or a dotted 6/8 rhythm such as the meter commonly found in Irish jigs—may also support smooth speech production for individuals with aphasia.

Due to the small sample size, it is also possible that differences in pause duration were due to external factors. For instance, the participant was practicing a different verb in both samples. It is possible that he simply found it easier to generate agent-patient pairs for one verb than for the other.

Zhang's MeVNeST sessions were based on improvised music. One of the strengths of this music therapy methods is its flexibility. However, this flexibility complicated the identification of strum conditions in this study. Zhang used a variety of rhythms across and within sessions. She also tended to adapt her playing to the participant's sung and spoken rhythms—pausing within participant pauses, for example, or simplifying a complex strum by switching to whole notes. The definitions for this study are therefore broad in order to apply to as many samples as possible. In a study devoted to the connection between accompaniment rhythm and speech pauses, it would be possible to establish more precise criteria and to hold more consistently to the predetermined rhythms during study sessions.

Future Research

This study raises potential opportunities for future research. Foremost is the possibility of examining the clinical use of syncopated rhythm in music therapy with IWA's. This study raises the possibility that there may be therapeutic benefit to the use of specific rhythms with this population. Future research may examine rhythms beyond the traditional 4/4 time, such as dotted 6/8 meter to see whether rhythms that closely mirror natural speech patterns may have a beneficial effect on client fluency.

The protocol developed in this study may also provide a foundation for further studies. In developing the protocol for this study, the author discovered the importance of consistency in coding guitar rhythms. Since the audio clips used for this study were drawn from improvisational MeVNeST sessions, many samples were unusable or provided only a weak example of a given strum condition. Future studies would be able to prioritize the use of a consistent, previously defined strum pattern that could be established early in the session.

Conclusion

This study sought to analyze the impact that strum pattern may have on IWA's fluency of speech. Future research into this topic should emphasize the practical use of rhythmic elements in music therapy with IWA's, particularly guitar rhythm, as this is a common accompaniment instrument in music therapy sessions. This knowledge would allow music therapists to better support their client's growth and progress towards their goals.

References

Boersma, P. and Weenink, D. (2021) Praat: doing phonetics by computer (6151) [Computer software].

Davis, G. A. (2011, February). *Aphasia therapy guide*. National Aphasia Association. Retrieved February 27, 2022, from https://www.aphasia.org/aphasia-resources/aphasia-therapy-guide/

Edmonds L. (2018) VNeST. In: Kreutzer J.S., DeLuca J., Caplan B. (eds) Encyclopedia of Clinical Neuropsychology. Springer, Cham. https://doi-org.libproxy.temple.edu/10.1007/978-3-319-57111-9_9130

Fujii, S., & Wan, C. Y. (2014). The role of rhythm in speech and language rehabilitation: The SEP hypothesis. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00777

Hobson, M. R. (2006). The collaboration of music therapy and speech-language pathology in the treatment of neurogenic communication disorders: Part I--diagnosis, therapist roles, and rationale for Music. *Music Therapy Perspectives*, *24*(2), 58–65. https://doi.org/10.1093/mtp/24.2.58

National Aphasia Association. (n.d.). *Home*. National Aphasia Association. Retrieved February 25, 2022, from https://www.aphasia.org/

Lee, Y. S., Thaut, C., & Santoni, C. (2018). Neurological Music Therapy for Speech and Language Rehabilitation. In M. Thaut (Ed.), *The Oxford Handbook of Music and the Brain* (pp. 714–737). essay, Oxford University Press. Retrieved February 24, 2022, from https://www-oxfordhandbooks-com.libproxy.temple.edu/view/10.1093/oxfordhb/9780198804123.001.0001/oxfordhb-9780198804123.

Lim, K.-B., Kim, Y.-K., Lee, H.-J., Yoo, J., Hwang, J. Y., Kim, J.-A., & Kim, S.-K. (2013). The therapeutic effect of neurologic music therapy and speech language therapy in post-stroke aphasic patients. *Annals of Rehabilitation Medicine*, *37*(4), 556–562. https://doi.org/10.5535/arm. 2013.37.4.556

Stahl, B., Kotz, S. A., Henseler, I., Turner, R., & Geyer, S. (2011). Rhythm in disguise: Why singing may not hold the key to recovery from aphasia. *Brain*, *134*(10), 3083–3093. https://doi.org/10.1093/brain/awr240

Thaut, M. H., & McIntosh, G. C. (2014). Neurologic music therapy in stroke rehabilitation. *Current Physical Medicine and Rehabilitation Reports*, *2*(2), 106–113. https://doi.org/10.1007/s40141-014-0049-y

Thaut, M. H., McIntosh, G. C., & Hoemberg, V. (2015). Neurobiological foundations of neurologic music therapy: Rhythmic entrainment and the motor system. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.01185

Zhang, J.-W. (in progress). *Developing Improvised Generative Speech Protocol for People with Aphasia: Music Enriched Verb Network Strengthening Treatment (MeVNeST)* (dissertation).

Zumbansen, A., Peretz, I., & Hébert, S. (2014). The combination of rhythm and pitch can account for the beneficial effect of melodic intonation therapy on connected speech improvements in Broca’s aphasia. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00592

Appendix

Praat Coding Key and Definitions

I. **Tier 1—Categories:** Coding of transcription data

a. Clinician responses:

i. **I**=scripted instruction or prompt

1. e.g., "Who chops what?"

ii. **S**=semantic cue

1. e.g., "Who might brush a horse?"

iii. **Ph**=phonetic cue

1. e.g., "Try a word that starts with *ss*."

iv. **Rf**=reflection of response

1. e.g., Participant says, "Barber brushes hair." and clinician replies, "The barber brushes the hair."

b. Participant:

i. **F**=filler word or sound

1. e.g., "umm"

ii. **A**=spoken or sung MeVNeST response

1. e.g., "Chef chops carrots"

iii. **M**=Musically licensed pause

1. e.g., waiting for downbeat of next phrase to sing response

c. Both:

i. **O**=other response

1. e.g., "I'm not sure"

2. e.g., "That's right"

ii. ******=silence

II. **Tier 2—Transcript:** What is being said and by whom

a. **P**=participant

b. **J**=first initial of music therapist

c. **G**=first initial of speech-language therapist

d. ****** = silence

III. **Tier 3—Music:** A running description of guitar strum pattern, usually without standard abbreviations

a. e.g., Guitar plays steady quarter-note downstrokes

b. **EE**=expected entrance

i. i.e., The participant's naturally sung entrance was at the beginning of the time stamp, and they did not come in

IV. **Tier 4—Notes:** Miscellaneous observations that could add to or clarify Tiers 1-3

V. **All tiers:**

a. **--** =continuation of the last entry

i. e.g., / downstroke / -- / -- / fingerpicking / -- / back to downstroke /

b. Time blocks that are left blank mean that the audio segment is not relevant to a particular tier