COMPARISON OF SPEECH RHYTHM BETWEEN SOLO READING AND CHORAL READING TASKS IN ADULTS WHO STUTTER

Deepthi, D.

Register Number: 14SLP010

A Dissertation Submitted in Part Fulfillment of Final Year

Master of Science (Speech Language Pathology)

University of Mysore, Mysore

ALL INDIA INSTITUTE OF SPEECH AND HEARING

MANASAGANGOTHRI, MYSORE-570 006

MAY, 2016

CERTIFICATE

This is to certify that this dissertation entitled **"Comparison of speech rhythm between solo reading and choral reading tasks in adults who stutter"** is a bonafide work submitted in part fulfilment for degree of Master of Science (Speech-Language Pathology) of the student Registration Number: 14SLP010. This has been carried out under the guidance of a faculty of this institute and has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysore
May, 2016

Dr. S.R. Savithri
Director
All India Institute of Speech and Hearing
Manasagangothri, Mysore-570006

CERTIFICATE

This is to certify that this dissertation entitled **"Comparison of speech rhythm between solo reading and choral reading tasks in adults who stutter"** has been prepared under my supervision and guidance. It is also been certified that this dissertation has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysore
May, 2016

Guide
Dr. Santosh M
Reader in Speech sciences
Department of Speech- Language sciences
All India Institute of Speech and Hearing
Manasagangothri, Mysore-570006

DECLARATION

This is to certify that this dissertation entitled **"Comparison of speech rhythm between solo reading and choral reading tasks in adults who stutter"** is the result of my own study under the guidance of Dr. Santosh M., Reader in Speech Sciences, Department of Speech – Language Sciences, All India Institute of Speech and Hearing, Mysore, and has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysore, **Registration No. 14SLP010**

May, 2016

ACKNOWLEDGEMENT

At the end of a long six year journey, I would first like to thank my family back home. Papa, Mamma and Deeksha. You guys have been my inspiration and my pillar of support. Thank you for teaching me to be what I am today. Without you'll I would not have survived.

A big Thank you to Santosh Sir! Sir, thank you for being my first mentor when everything here was new and daunting. Thank you for your, patience, wisdom, and enthusiasm. You provided me with all that I needed to successfully complete this dissertation. Thank you for tolerating our mistakes during the JC and being a great guide to work with! It was a pleasure working under your guidance!

Thank you, Dr. Savithri, for giving us the opportunity and resources to complete this dissertation. I thank everyone in DCS, especially Dr. Sreedevi for giving me the permission to use the clinical resources.

My heartfelt gratitude to all the patients who spared their time and willingly participated in this study, and all my juniors and classmates who helped with the data collection.

It is always the people behind the curtain who work the hardest to put up a show. Nike sir, you have worked harder than I did to help with the rhythm analysis. Without you this dissertation would not have been completed. Thank you for your patience, time and effort. I'm immensely grateful to you for all your help and your valuable input.

Dhatri Ma’am, thank you for letting me hijack your recorder! Aparna Ma’am, Abhishek Sir. Thank you for helping me begin my dissertation and giving me the resources I needed. Girish Sir, thank you for being the voice behind my dissertation!

Rashika. If not for you constantly panicking I would have forgotten to complete my work. Thank you!

Janani, my baby. Thank you for being my daily dose of entertainment. I hope all your dreams come true! Anu, Akshu, Ayesha begum, Nivu, Nikki. My family in AIISH. Thank you for being my lifeline here. Adi and Mani. My crazy counterpart and my soul sister. Thank you for being there with me every step of the way!

Aiyappa. Thank you for bearing with me. You made the last few months easier. I wish you all the happiness in the world!

Last but not the least, I am eternally grateful to ISH and AIISH for laying the foundation I need for my career. I thank all my teachers for guiding me throughout my life.

Table of contents

List of Tables

List of Figures

CHAPTER 1

Introduction

Among the many fluency inducing conditions, choral reading induces maximum fluency (Andrews, Howie, Dozsa, & Guitar, 1982; Cherry & Sayer, 1956). Majority of the studies have reported the reduction in fluency at around 90-100%. Despite such significant reductions in the dysfluencies, valid explanations for its effectiveness are not known. A few notions put forth by the previous researchers are, a decrease in the communicative expectations (Eisenson & Wells, 1942), introduction of new vocal patterns (Wingate, 1976), presence of an external timing mechanism (Johnson & Rosen, 1937), unusual speech pattern (Ingham & Packman, 1979) and as an act of fluent imitation (Kalinowski & Saltuklaroglu, 2003; Skoyles, 1998). Wingate (1976) suggested that during choral reading, continuity of speech is achieved by emphasizing on vocalisation which is triggered by the prompts one receives from the second speaker. Adams and Ramig (1980) studied the vocal characteristics on choral speech in stutterers, and observed that vowel durations were altered by the stutterers during the choral reading condition, where in the stutterers showed reduced vowel duration when compared to the controls.

Kiefte and Armson (2007) attributed that in choral speech the subject has to pace with a second speaker and this influences a change in stuttering frequency. Thus, the discriminating feature of choral reading is identified as an *intoning* pattern or change in *rhythm* that influences one's speech. Therefore, the question arises, whether the dysfluencies reduced during choral speech is due to alterations in *rhythm*. However, till today, as per our knowledge, there is no published study comparing the speech rhythm characteristics between solo reading and choral reading tasks in persons who stutter. Thus, the aim of the present study is to investigate whether there is any change in speech rhythm during choral reading when compared to solo reading condition in Kannada speaking adults who stutter.

Rhythm in speech is made of multiple ‘strong’ and ‘weak’ components (Schane, 1979). In the past, to study changes in speech rhythm multiple duration-based metrics have been used (Dellwo, 2006; Ling, Grabe, & Nolan, 2000; Ramus, Nespor, & Mehler, 1999). Typically these metrics are constructed from the vocalic and intervocalic intervals of speech sounds to acoustically break down speech. Pair-wise variability indices (PVIs) are based on the premise of vocalic and consonantal interval durations. However, these methods try to assess rhythm in its syntagmatic form by finding differences between intervals of vocalic sounds or consonants (Grabe, & Low, 2002). Their applications are reduced in disordered speech as the clear boundaries are not available for the precise marking of vocalic and consonantal intervals. A relatively newer concept of rhythm metrics is the Envelope modulation Spectrum (EMS) (Drullman, Festen, & Plomp, 1994; Greenberg, Arai, & Silipo, 1998).

Traditionally, the stress patterns of speech (strong and weak syllables) were thought to be cued predominantly by differences in a speaker’s fundamental frequency (Fry, 1955, 1958). However, more recent studies have found that amplitude and duration cues aid in better discrimination of prosodic stress than fundamental frequency (Greenberg, 1999; Kochanski, Grabe, Coleman, & Rosner, 2005). Leong, Stone, Turner, and Goswami (2014) investigated the rhythm contribution of slow amplitude modulation (AM) cues that are present within the envelope of speech (i.e., up to 50 Hz) rather than on faster temporal modulations (i.e., 50–500 Hz; Rosen,

1992) that typically carry information about fundamental frequency as well as possibly some information about syllable onsets.

The EMS is a representation of the slow amplitude modulations in a signal (Liss, LeGendre, & Lotto, 2010). Steeneken and Houtgast (1980) first used this spectrum to quantify effects of room acoustics on speech. The EMS is said to be an acoustic measure that does not consider the signal's linguistic make up. It represents the power of a spectrum of the signal's amplitude envelope and the amplitude variations across frequencies. EMS uses a 'modulation spectrum' that is ideally between 3-5 Hz (Goswami & Leong, 2014; Greenberg, Arai, Grant, 2006; Houtgast & Steeneken, 1985; Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995) such that it provides information regarding rhythm of speech, regardless of the rate of speech. The average syllable duration is calculated to be 200 ms, so using modulation of amplitudes around 5 Hz are expected to reveal information about patterning of the syllables in speech (Ahissar, Nagarajan, Protopapas, Mahncke, & Merzenich, 2001; Greenberg, Carvey, Hitchcock, & Chang, 2003; Luo & Poeppel, 2007). Prosody and information about stress can be obtained by implementing slower amplitude modulations i.e., slower than syllable rates (Ghitza & Greenberg, 2009). On the other hand, faster rates, of about 50 Hz will indicate linguistic information regarding articulation, voicing parameters, and vowel identification (Rosen, 1992).

In the past few studies have used EMS to characterize rhythm. Tilsen and Johnson (2008) focused on the 700 Hz to 1300 Hz region to determine the perceptual units of stress in English. This specific frequency region is said to help examine the vowel nuclei, bursts, voicing, fricatives to understand rhythmic patterns in disordered speech. Envelope modulation spectra was used by Liss, et al., (2010) to discriminate between dysarthria types using six parameters defining rhythm. As EMS is an automated technique and it also has the capacity to manage the non-linguistic parts of speech such as pauses, noise, without altering the analysis, it is more suitable while analyzing disordered speech. (Liss et. al.2010).

Current study adopted AM Phase Hierarchy (AMPH) model proposed by Leong (2013). The essential premise of the model is that prosodic rhythm patterns in speech may be inferred from the phase relationships between hierarchically-nested slow amplitude modulations (AMs) in the acoustic signal. Following a method of Principal component analysis of 24 channels of speech spectra, 3 amplitude modulation rates were identified as being crucial in the analysis of speech rhythm. They are provided in Table 1 (Leong, 2013).

Table 1
Modulation rate bands identified from component loading patterns.

Linguistic unit	**Modulation rate band (Hz)**
Prosodic stress (Modulation band 1)	0.9-2.5
Syllable (Modulation band 2)	2.5-12
Phoneme (Modulation band 3)	12-40

The purpose of this study is to investigate whether the reduction of dysfluencies observed during choral reading is due to alterations in the rhythm in the speech (measured through changes in envelope modulation/ amplitude modulation) of persons with stuttering.

The objectives of the study are as follows:

1. To compare the change in percentage of dysfluencies (%SS) between solo reading and choral reading.
2. To compare speech rhythm (using amplitude modulation/ envelope modulation spectra) between solo reading and choral reading tasks in adults who stutter.
3. To identify the rate of amplitude modulation (stress, syllable or phoneme) at which the speech rhythm varies to a greater extent between solo reading and choral reading.

CHAPTER 2

Review of literature

It has been well-documented by numerous studies that when persons who stutter (PWS) are involved in 'speaking in unison' with a second speaker with normal fluency, they are able to produce natural and effortless speech and this phenomenon is termed as choral speech. (Bloodstein & Bernstein-Ratner, 2008; Kalinowski & Saltuklaroglu, 2005; Silverman, 2003). It has been found that as PWS indulge in choral reading either with memorized passages or have to read prepared texts in close approximation with a second speaker, their stuttering behaviours declines by almost 90-100%. Studies on choral reading date back to as early as 1930's.

Johnson and Rosen (1937) were one of the first few researchers to confirm the "choral speech effect" via their experiments on PWS. Eighteen adult stutterers participated in their study wherein they were made to read three different 500-word passages using 12 different imposed speech patterns such as speaking slowly, speaking fast, whispering, speaking with low intensity, high intensity, singing, metronome arm swing, sing song, high pitch, chorus stutterer, and chorus normal. Six of the twelve conditions demonstrated no stuttering, namely, singing (0%), metronome (.06%), arm-swing (.01%), sing-song (.09%), reading in chorus with another stutterer (0.01%), and reading in chorus with a non-stutterer (0.01%). It was suggested that the conditions in which maximum reduction of stuttering occurred were those that altered the speech pattern of PWS with an imposed and definite rhythm, which included choral reading as well.

Barber (1939) conducted an experiment on 18 persons with stuttering (3 male and 15 female) whose ages ranged between 15 – 33 years (Mean= 23.8years). Each subject read 14 different 500-word passages under 14 altered conditions of choral reading, as given below in Table 2. It was noted that stuttering was almost eliminated in conditions 2-5 where the subject read the same material with an accompanist. The mean reductions in percentage of syllables stuttered, across the second to fifth condition, were 97.8% (S.D=5.6), 95.1% (S.D=12.0), 92.7% (12.2), and 90.7% (S.D=22.3) respectively.

Table 2

Experimental conditions (Barber, 1939).

Sl. no	**Condition**
1	Subject reads alone with no experimental distraction operating
2	Two stutterers and one normal speaker read with subject, the same material being read by all.
3	Two stutterers read same material with subject.
4	Normal speaker reads same material with subject.
5	One stutterer reads same material with subject.
6	Two stutterers read a passage different from that being read by subject.
7	Normal speaker reads a passage different from that being read by subject.
8	One stutterer reads a passage different from that being read by subject.
9	Normal speaker reads nonsense syllables while subject reads a meaningful passage.
10	One stutterer reads nonsense syllables while subject reads a meaningful passage.

11	Normal speaker phonates "ah" at relatively constant pitch while subject reads a meaningful passage.
12	Phonograph record of same normal speaker phonating "ah" at approximately same pitch is played while subject reads.
13	Unpatterned mechanical noise is made by normal speaker while subject reads.
14	Subject reads alone with no experimental distraction operating.

Thus, it was noted in this study that the most marked difference in the stuttering percentage occurred in the condition matching choral speech most closely. However, an unavoidable limitation of the condition is that the linguistic material used in united reading must be matching to produce the desirable inhibitory effects. Therefore, majority of the studies have all used prepared texts in investigating the choral speech effect. An interesting observation has been that if the extraneous speech signal present in choral speech is removed, stuttering recurs almost instantly. It is also argued that though it is termed choral speech, it is said not to be entirely synchronous in nature owing to the problem that it is impossible for two speakers to speak in direct agreement with each other. Despite the temporal asynchrony choral speech has earned the label of being the gold standard for natural easy speech in PWS where they require no training.

A large number of theories have been proposed regarding mechanisms of stuttering inhibition under choral speech. The "reduction of communication load, distraction, rhythm alterations, masking and altered vocalizations" has been attributed for the reduced dysfluencies in PWS. Studies in the 1940's have been done to determine the reasons for the reduction in stuttering in the choral speech condition. In one of the initial studies, Eisenson and Wells (1942) investigated the influence of communicative responsibility on the speech performance of PWS in choral speech situation, where they questioned whether the absence of responsibility during communication, facilitates fluent speech production. Nineteen male stutterers participated in the study. They were asked to read a passage in two choral reading conditions: 1) choral reading without communicative responsibility where they were first asked to read in unison with another examiner, 2) reading into a microphone in unison, while they were being heard by another room with a group of listeners simulating a radio situation contributing to the communicative responsibility situation. The results revealed that the stutterers performed poorly in the 2^{nd} situation (with an average increase in stuttering percentage of 60.2%) and suggested that communicative pressure exerted on the PWS during speaking alone were reduced during the activity of speaking under choral speech with only the second speaker being present but not when they were being heard by other people.

In Pattie and Knight's (1944) study, they denied the explanation by Eisenson and Wells (1942) by recruiting twelve stutterers who read before a small audience under several conditions: a) alone, in unison with a second person, b) in unison with a person whose voice was heard by telephone, and (3) in unison reading of different passages. The number of stuttered blocks showed no consistent influence of reading in unison but showed a marked influence of hearing unison reading of the same passage. Thus, they suggested that the benefits of choral reading depend upon more factors than just the decrease of conspicuousness and emotional tension.

In another study done by Bloodstein (1950), responses from 204 questionnaires filled by PWS, regarding fluency-enhancing conditions were evaluated. Questions were also related to the choral speech effect. Of the participants, 95% reported almost complete eradication of their stuttering while using the choral speech condition when same text reading was carried out, whereas only 13 reported similar reductions with 25 different texts. Cherry and Sayers (1956) tested the above findings where they conducted an experiment, and also reported similar levels of stuttering reduction during reading of similar texts and even when the second speaker switched texts. The results showed that the same could not be achieved when the speakers started out with different texts. A reduction in stuttering was also noted when nonsense syllables were produced by the normal speaker instead of a text and when an audio recording was played in reverse while the subject read the text from beginning to end.

Wingate (1969, 1970) was one of the first few authors to suggest the possible reasons of great reduction in stuttering in choral reading in addition to other fluency inducing conditions such as imposed rhythm, shadowing, singing, and reduced auditory sensitivity. He posited his hypothesis of *changed phonation* after reviewing evidence for reduced stuttering and changed speech patterns during the above mentioned conditions.

Adams and Reis (1971) hypothesized that the frequency with which vocalization must be initiated in a particular speech segment and the frequency of attendant stuttering are positively related and suggested that the second statement must be true. They included 12 male and 2 female adults with stuttering (Age range of 15-26; mean age= 19.57). Two 100 word passages with one having predominantly voiced sounds at the beginning of the word and another with a combination of voiced and unvoiced, were used as reading material. Each subject was asked to read the passage aloud 5 times in two sittings and the frequency counts were tallied for each reading. The results of the study showed that significantly less stuttering occurred in the all voiced passage (P=>0.03). The authors thus interpreted their finding of reduced stuttering in totally voiced reading as due to the reduced need for phonatory adjustments.

A later claim by Wingate (1976) is that at least one thing stutterers do during chorus reading is slow their speech rate. Wingate concluded that "chorus-reading is clearly attended by a reduction in individual speech rate". Ingham and Carroll (1977) studied the change in stuttering frequency in solo reading and choral reading, and also the speech rate of nine stutterers by listener judgement tasks wherein speech language pathologists were asked to decide whether speech samples belonged to any one of the two following conditions i.e., ; 1) from different reading situations (that is, one made while reading alone and one made while simultaneous reading was occurring), 2) or from the same reading situations (that is, both reading alone or both reading simultaneously). The results revealed that chorus reading condition was associated with a substantial reduction in stuttering. The group's mean speech rate was found to have a significant increase during chorus reading. The listener judgement task revealed an accuracy of 57%.

Adams and Ramig (1980) conducted a study on 10 PWS, ranging in age from 15 to 40 years and with a mean age of 26 years 11 months (9 males and 1 female). The severity of their stuttering ranged from mild to severe. A group comprising 1 female and 9 male normal speakers was considered as the control group (aged 21 to

31; Mean=25 years). A 500 syllable prose passage with neutral content was used in this study. Two conditions were introduced in the study: 1) in the control condition, subjects were given instructions to read the experimental passage aloud in their habitual rate and manner; 2) in the experimental condition, subjects were told to read the same passage aloud and in unison with a normal speaker's tape recorded reading of the same prose. The audio tape was used in order to avoid the influence of the subject’s speech on the model speaker. The authors hypothesized that the normal speakers and PWS might emphasize vocalization by increasing their vowel durations and vocal SPL, and by sustaining phonation across word boundaries, thus four dependant variables were considered, namely, subjects' dysfluency, vowel duration, peak vocal SPL and percent continuity of phonation scores. The results revealed that subjects exhibited significant reduction in dysfluencies in choral reading condition ($p = 0.047$) and also a significant decrease in the vowel durations between control and choral conditions.

Andrews Howie, Dozsa, and Guitar (1982) studied the effect of 15 fluency-inducing conditions including choral speech on 3 adult males with stuttering. The experiment was conducted on 6 baseline trials and 15 experimental trials (one for each condition). For the choral speech condition the subjects were asked to read in unison with examiner from a newspaper. The first 300 syllables of the speech samples were analysed to measure stuttering frequency (%SS). The authors found that stuttering reduced in all conditions except in three conditions (speak and write, relaxed and alone with cards). Choral reading resulted in maximum decrease in the frequency of stuttering (0% SS, Mean percentage decrease in stuttering=100%) and there was also no change in speech rate. The absence of change in fluent speech rate during chorus reading supports Ingham and Packman's (1979) finding that decreased fluent speech rate is not necessary for stuttering reduction under this condition. Authors also suggest the reduced load of language formulation to aid in fluent speech. At the linguistic level longer phonation durations would reduce the frequency of decisions about what is to be said; at the prosodic level longer phonation durations would reduce the frequency of control decisions concerning intensity, intonation, and pausing.

A more novel explanation to the choral speech effect was provided by Bloodstein (1995) where he credited the effect to 'distractions' in the "distraction hypothesis". The hypothesis states that 'any condition that reduces stuttering is only a distracter'. Webster and Lubker (1968) mentioned that rhythm, masking noise, choral reading , prolongation of speech sounds, whispering, delayed auditory feedback tend to reduce stuttering and improves fluency in PWS. They examined choral reading in light of the "Auditory interference theory" (AIT). They were under the impression that rhythmic cueing stimuli or prediction cueing stimuli like experienced in choral reading, when viewed in terms of AIT, provide sensory stimulation which enables the stutterer to produce speech output signals that have little or no interference from the returning auditory signal. Apparently, the cueing stimulus governs the initial release rate of syllables to be produced by the stutterer thus removing the rate control function from guidance by auditory feedback. AIT leads to a suggestion that the cueing properties of rhythmic stimuli permit speech output control signals to override any interference that may occur from auditory feedback. While the above studies have focused on both behavioural and psychological aspects of stuttering, Armson and Kiefte (2008) suggested that choral speech generates an "external timing, or pacing rhythm". However the notion is questionable because there is no evidence of the speakers training their timing during choral reading and the signals are also asynchronous. In addition, fluency has been observed to be maintained if the second speaker changes passages during the reading and comparable inhibitory effects are observed under sustained /a/ production, backwards and forward flowing stuttered speech (Wingate, 1976; Dayalu, 2004).

Choral speech has a "magical" effect on stuttering of successfully removing nearly all, 90-100% of overt stuttering behaviours (Andrews et al, 1982; Cherry & Sayers, 1956; Kalinowski, Guntupalli, Stuart & Saltuklaroglu, 2004). This has been attributed to the presence of a "second speech signal" that arises due to the second speaker joining in, speaking in a normal manner. The signal may either be acoustic or visual, and must contain information about the gestural status of the vocal tract.

Kalinowski, Stuart, Ratstatter, Snyder, and Dayalu (2000) introduced the novel phenomenon of visual choral speech, which is described as the fluency enhancement occurring via visual gestures of speech when there is no auditory feedback. Ten participants (8 males 2 females; Mean age=27.9, SD=9.4) with moderate to severe stuttering participated in the study. They were asked to memorize sentences presented via cards and recited the sentences under two conditions (1) Non Visual Choral Speech (NVCS) condition – without the external speaker; (2) Visual Choral Speech (VCS) – where the subject was asked to recite while focusing on the articulatory movements of an external speaker who silently mouthed the words. Mean values for stuttering frequency for NVCS and VCS conditions were 77.2 (Standard error= 16.5) and 16.5 (Standard error= 6.6) respectively. It was noted that stuttering reduced by 80% in the VCS condition.

Kalinowski, Dayalu, Stuart, Ratstatter, and Rami (2000), examined the influence of a stuttered speech signal that was externally provided to the subject and measured the effects on stuttering frequency. It was compared to an externally generated normal speech signal. They also aimed at examining the specific components of the “second speech signal”, which might be contributing for the encouragement of fluency in people who stutter. The authors used reading material that were both congruent and incongruent in order to identify whether the material being read has a role to play in the inducement of fluency. Ten individuals with stuttering (eight males, two females, Mean age=27.9, SD=9.4) whose severity ranged from moderate to severe were recruited for the study. Two experiments were conducted in the study wherein the first experiment involved the subjects reading with an external speech signal that was incongruent fluent or stuttered speech. Incongruent fluent speech was when the external signal was fluent speech but a different reading passage and incongruent stuttered speech was when the external signal was dysfluent speech of a dissimilar reading passage. In the second experiment subjects listened to four continuous speech signals : a steady state neutral vowel /a/, a three vowel train representing the three corners of the vowel triangle /a-i-u/, a steady state consonant /s/, and a three consonant train /s-sh-f/. The subjects also read a control passage which served as the non altered auditory feedback (NAF).The samples were analysed for changes in stuttering frequency. The results indicated that greatest reduction in stuttering was noted in the fluent continuous condition where the subjects read in unison with a fluent external speaker. Mean values for NAF were 35.4% SD=25.11 whereas in the fluent condition mean values of stuttering frequency were found to be 8.6%, SD= 13.2.

Rami and Diederich (2005) conducted a study on 8 adults aged between 19-54 years to explore the effects of choral reading, reverse speech at normal speed, reverse speech at half the speed and no auditory feedback. They reported that the choral reading condition tested here produced a statistically significant decrease in the frequency of stuttering. Their results supported Pattie and Knight (1944) claim that the fluent reader behaves as a pace-setting mechanism for the individual with stuttering. When they read in harmony with another person, it is said to impose "*a steady rhythm of speech"* that helps them talk more fluently. It was suggested that the third stimulus, which was reversed speech at half speed, probably contained intonation patterns comparable to everyday speech, which could have acted as a pace-setter for the participants.

Rami, Kalinowski, Ratstatter, Holbert, and Allen (2005) attempted to investigate the use of choral reading with a more novel stimulus where filtered components of speech and whispered speech were used as the choral reading stimuli. They measured the effects of the altered form of the external signal on the frequency of stuttering on 12 adults with stuttering (10 men and 2 women, age range= 19.2-53.1 years; mean=37.7 years). Three passages of 300 syllables, were read by a normal adult male (aged 28 years) and the samples were low-pass filtered with centre frequencies at 100 Hz (approximate glottal source), 500 Hz (source and first formant), and 1 kHz (source and the first two formants), (to arrive at three models for choral reading), along with a whispered passage, a normal passage and a control condition (non-altered auditory feedback). The authors postulated that the control condition of non-altered feedback would not produce fluent speech. The frequency of stuttering was analysed in each of the conditions. The choral speech effect was found to be replicated in the study by a marked reduction in the stuttering frequency in choral reading condition (mean =3.6%, S.D= 2.9) along with whispered speech (Mean=4.1, S.D=3.6) and filter-500 (mean= 6.6%, S.D=10.4), filtered-1000 (mean= 2.5, S.D= 3.4), as compared to non altered auditory feedback condition (mean=24.5%, S.D=35.5). The authors concluded by suggesting that the filter characteristics of the vocal tract may be the cues that help in decreasing stuttering in choral reading condition.

Kiefte and Armson (2007) compared the effects of choral speech and altered auditory feedback (AAF) on stuttering frequency in order to identify the underlying aspects of choral reading which contributed to the reduction in stuttering. 17 adults with stuttering, aged between 19-57 years (Mean= 34), participated in the study. They were subjected to two conditions 1) choral reading with live speaker, 2) delayed choral speech which was delivered to the participant with a delay of 60 ms. Four baseline and six experimental conditions were presented including solo reading, choral reading (live accompanist), choral reading (delayed live accompanist), choral reading (recorded accompanist), choral reading (recording of participant as accompanist and altered auditory feedback (AAF). Results of the study showed marked reduction in stuttering in choral reading and AAF conditions when compared to solo reading and the reduction in choral reading condition was most dramatic. The manipulated choral reading experiments resulted in greater reduction in stuttering possibly because of the external pacing that is induced during reading with an accompanist.

Balasubramanian, Cronin, and Max (2009), used three conditions in their study, a) repeated readings of the same material, (b) choral readings with a second speaker, and (c) readings with delayed or frequency-altered auditory feedback. Two individuals (1 male aged 61 years, 1 female aged 60 years) with acquired Neurogenic stuttering post cerebrovascular accident were taken as subjects. As elaborated in the study, both patients showed a decrease in the stuttering frequency in the choral reading condition (50%). However the reduction was not as marked as that observed in persons with developmental stuttering who show 90% reduction in stuttering (Ingham & Packman, 1979). The figure below shows the stuttering frequency of one subject across each of the experimental conditions. It was also noted that the reduction in dysfluencies in the choral condition was more consistent (4/5 readings) when compared to the other two conditions.

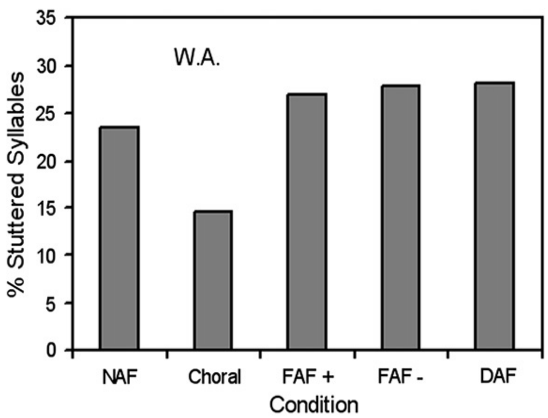

Figure 1: Dysfluency levels of female patient across experimental conditions.

Hudock (2012) conducted a study on 11 participants who were nine males and two females with a mean age of 30.2 years (*SE* = 4.9, range 18 – 72). Four non standardized passages were used as reading material and normally fluent female individuals served as the external reader. Three conditions were used in the experiment 1) choral speech, 2) shadow speech with the person who stutters maintaining the lead speaker position, 3) shadow speech with the person who stutters maintaining the lag speaker position, and 4) baseline. The results revealed that stuttering frequency significantly reduced during all three experimental conditions.

Stuttering was repressed by 95% in the choral speech condition. However, in the lead and lag speaking conditions significant differences were not observed. (*i.e.*, 79% and 81% respectively). The stuttering frequencies being similar during both, lead and lag conditions disprove notions proposed by authors of 'external timekeeper models' and 'demand and capacity models' which attempt to explain alleviation of stuttering during the perception of second signals. They explain that, the lead speaker conditions, requires generation and self monitoring of self-produced rhythm and pace, and increasing cognitive demands for language formulation tasks. Models highlighting the pacing and rhythm aspects suggest that perceiving the second signals externally helps rely on the external pacing signals that in turn reduce stuttering.

Park and Logan (2015) studied the extent to which the effect of choral speech depended on the intactness of temporal speech cues. Eight adults with stuttering (7 males, 1 female; M=22 years, SD=12.27 years) and eight adults without stuttering (6 males, 2 females; M=23 years, SD=13.75) participated in their study. They were asked to read 20 sentences in three different conditions, to provide a total of 60 sentences. The experiment included a solo reading condition and three other conditions namely 1) typical choral speech, 2) altered choral speech and 3) multi-speaker babble as choral signal. As reported by the authors, significant differences were present in stuttering frequency (p=0.012). Persons with stuttering were more fluent under the choral speaking conditions than the solo speaking condition. However it was stated that the results obtained do not support views of the temporal modelling or the articulatory gesture modelling hypotheses, because fluency enhancement was observed during the 1) temporally altered choral speech condition, wherein the temporal information of speech was altered, to create a mismatch with normal prosody, 2) the conversational babble condition, and the 3) typical choral speech condition. Hicock, Houde, & Rong (2011) suggested that in choral speech, the speech-related information is readily available, which enables the auditory system to match auditory targets in the speech motor plan with the somatosensory output which indirectly aids in compensating a deficit in the sensorimotor integration circuit.

Mirror systems and Fluency enhancement

The concept of mirror neurons which have been found in the premotor cortex of monkeys, explains the phenomenon of imitation and its effects on learning. Efforts have been made by authors Kalinowski and Saltukluroglu (2003) to explain the ameliorative effects of choral speech as a result of the functioning of mirror neurons. The use of choral speech or imitated speech has been said to override the central involuntary dysfluency (blocks) as mirror systems were present innately, before the onset of stuttering. In other words, gestural imitation of fluent speech enhances fluent speech and inhibits stuttering. However the inhibition only lasts as long as the choral speech signal and it relapses to previous levels when it is terminated. Thus choral speech is said to be a form of direct imitation and it maintains the strong fluency enhancing capabilities, suggesting there is some temporal flexibility in the system itself. According to Saltukluroglu and Kalinowski (2006), during choral speech a gestural mirror is provided which inhibit the neural block and functions via 'direct and immediate transposition of speech perception to production'.

Amidst multiple explanations with respect to the effects of choral speech on stuttering, many researchers regard stuttering to be a consequence of timing problems in speech production and perception (Van Riper, 1982). Similarly, Natke, Gosser & Kalveram (2001) regards stuttering as the result of a problem in rhythmical processing: According to them the subdivision of the speech signals into consecutive syllables, accelerates and automatizes speech production without the need of auditory feedback. They believed this mechanism to be impaired in people who stutter. Another fact that indicates the link between rhythmical processing is the circumstance that in performance styles where the timing is controlled independent of speech rate, like singing or whispering, stutterers tend to have fewer problems, (Glover, Kalinowski, Ratstatter & Stuart, 1996). Methods of timing control, like an artificial reduction of articulation rate, accompanying tapping or the use of a metronome are often used in stuttering therapy (Ptok, Natke & Oertle, 2006).

As explained above, several factors have been thought to contribute to the ameliorative effects of choral reading, the closest being the effect of a change in the rhythm of PWS and the distraction hypothesis. Traditionally the analysis of rhythm has been conducted using methods to obtain amplitude and duration cues, which have been found to contribute more strongly than an analysis of the fundamental frequency in prosodically prominent regions, and these have been extended to the analysis of speech rhythm more recently (Greenberg, 1999; Loukina, Kochanski, Rosner, Keane & Shih, 2011). Among the methods commonly used to describe and measure speech rhythm, they have been broadly classified as being either ‘duration-based’ or ‘amplitude-based’ approaches. The duration-based approach is identified by “rhythm-metric” measures. These are summary statistics designed to distinguish between languages with different perceived rhythmic qualities and incorporates the measures of duration with respect to consonantal (C) and vocalic (V) intervals. Consequently they enable the classification of languages as being ‘stress-timed’ versus ‘syllable-timed’ languages (Abercrombie, 1967; Pike, 1945). Commonly used indices include %V, ΔV, ΔC (Ramus et al., 1999) and these measure help quantify the relative proportions of vowels (vocalic intervals) and the standard deviation of vocalic and consonantal (interconsonatal/ intervocalic) durations in speech, whereas the measure of pair-wise variability indices (PVIs), (Grabe & Low, 2002) and rate-normalized measures like VarcoV and VarcoC (Dellwo and Wagner, 2003) focus on the relative variability in the length of successive consonantal and vocalic intervals.

In the field of acoustic studies, the changes occurring in amplitude, duration and frequency are considered to be prominent cues in prosodic rhythm (Hirst, 2009). The amplitude-based cues to rhythm have been identified to lie within the slow-varying "amplitude envelope" of speech. Within the speech envelope the moment of occurrence of a sound across frequency is termed as a "perceptual (P) center" and is cued by the slow-varying amplitude patterns in the signal. (Allen, 1972; Morton, Marcus, Frankish 1976; Scott, 1993, 1998; Villing, 2010). Thus the P-center is said to form the basis for the purposeful rhythmic timing of speech and for synchronization of speech between speakers (Cummins and Port, 1998; Cummins, 2003). The envelope onset time that is the moment at which the envelope begins controls/ influences the P-center and it is related perceptually to the envelope onset rise time. This synchronisation is the premise of choral reading.

In the arena of signal processing, the speech signal can be viewed as the product of a very fast varying carrier, which contributes to the fine aspects/ structure of speech and a considerably slower varying amplitude envelope. The slow varying amplitude envelope interacts with the fast varying carrier to modulate the amplitude of the carrier. The demodulation or breaking down of this envelope carrier conveniently aids in isolating the amplitude based cues to the P- centres and the subtle prosodic changes in rhythm patterns found in the original speech signal prior to demodulation. The amplitude envelope of the original signal itself contains multiple rates of amplitude modulation (AM), which forms a ‘modulation spectrum’ of a number of modulation rates, out of which not all are equally important for transmitting information about rhythm in speech. The 3-5 Hz range in the modulation spectrum has been identified to have the strongest modulation and is independent of the language or speech rate (Shannon et al., 1995; Houtgast & Steeneken, 1985; Greenberg et al., 2003; Greenberg, 2006).The average duration of a syllable in speech is approximately 200 ms, and the amplitude modulations around 5 Hz have been defined to relate to syllable-pattern information in speech (Greenberg et al., 2003; Ahissar et al., 2001; Luo and Poeppel, 2007). Amplitude modulations which are slower than the syllable rate i.e.; less than the 3-5 Hz, are said to be related to prosodic stress patterns (Greenberg et al., 2003; Ghitza and Greenberg, 2009). Rosen (1992) exposed that faster amplitude modulations up to 50 Hz contain linguistic cues to phonetic manner of articulation, voicing, and vowel identity). The role and contribution of different AM rates within the speech envelope is usually investigated with respect to speech intelligibility.

Similarly, Liss and Lotto (2010) adopted an automated analysis of the speech envelope modulation spectra (EMS) which aims at quantifying the rhythm in speech by splitting it within specific frequency bands. They aimed to study whether similar results as those in Liss et al (2009) study could be achieved to differentiate the Dysarthria types based on prosodic/ rhythm changes. In the EMS they include slow rate amplitude modulation spectra (up to 10 Hz) for 7 octave bands and six variables related to the amplitude and frequency are calculated. Preliminary studies have focused on the 700 to 1300 Hz regions as they are said to correspond to linguistic rhythm. However, according to the EMS the region between 125 and 8000 Hz allows the examination of rhythmic patterns arising due to vowel nuclei, voice.ng, bursts and fricatives among many others. The analysis provides six variables in eight frequency bands providing a total of 48 variables. In this study the authors recruited 43 speakers with Dysarthria, and each subject was asked to read five sentences. The modulation spectrum was obtained for each of the sentences using the EMS program to provide 48 variables. A discriminant function analysis was conducted on the 48 variables to arrive at the predictor variables which enable distinguishing Dysarthria types based on rhythm of speech. Each Dysarthria type revealed a set of variables which best predicts the type of Dysarthria based on the EMS analysis. The authors suggest that instead of a manual manipulation of the speech segments, identifying a set of predictor variables from the amplitude spectra would contribute to a more detailed and extensive analysis of rhythm.

As the previous literature has focused on the use intelligibility cues from amplitude envelope, recent research advances towards the *rhythm cues* from the speech amplitude envelope (Leong, 2013). These envelope-based rhythm cues help to support rhythmic synchronisation between speakers even when speech is unintelligible (Cummins, 2009). A novel approach in the automatized approach to rhythm analysis is the Spectral Amplitude Modulation Phase Hierarchy (S-AMPH) model (Leong, 2013) which is implemented here. This is said to be a low-dimensional representation of the speech envelope to use as the spectro-temporal representation underlying the new S-AMPH model. To achieve this, PCA analyses were applied in the spectral and modulation rate domains. Component loading patterns were derived and analysed for evidence of channel 'clustering' (i.e. peaks) and boundaries indicating the transition between different spectral/modulation bands (i.e. troughs). Based on these analyses, 5 spectral bands and 3 modulation rate bands were identified.

In the S-AMPH model, an AM hierarchy consisting of nested modulation patterns at 'Stress' (0.9–2.5 Hz), 'Syllable' (2.5–12 Hz) and 'Phoneme' (12–40 Hz) rates is extracted from the speech amplitude envelope for 5 octave bands (1) 100-300 Hz; (2) 300-700 Hz; (3) 700-1750 Hz; (4) 1750-3900 Hz; and (5) 3900-7250 Hz.. These 3 AM tiers represent the dominant non-redundant modulation structure that is present in the speech envelope at 3 different (but simultaneous) time scales. Leong (2012) determined the existence of these AM tiers via principal component analysis of a multi-speaker corpus.

The S-AMPH has been used in several studies involving both perception and production tasks of rhythm. Leong et al. (2013) conducted a rhythm judgement task where the goal of the experiment was to better understand how AM rates (i.e., Stress and Syllable) from the speech envelope contribute to listeners' perception of speech rhythm. Twenty-three adults (7 male; mean age 26.0 yrs, range 22.0–37.5 yrs) participated in the study as listeners. They had to identify the type of rhythm/ stress pattern (trochaic/ iambic) of amplitude modulated tone vocoded nursery rhyme sentences. The results showed that the non phase shifted data indicated that the Stress and Syllable AM condition transmitted the most rhythm pattern information, since participants were statistically the most accurate at making rhythm discriminations when presented with this AM combination (57.4% accuracy). Accordingly, it is concluded that the rhythm information contained within the envelope modulation spectrum is primarily located at Stress (2 Hz) and Syllable (4 Hz) rates, and that the perception of (English) speech rhythm depends in part on the phase relationship between these two key rates of AM.

Thus, it is possible to extract amplitude modulation information from the desired signal by obtaining the envelope modulation spectrum to identify the dominant rhythm pattern in the spoken utterance. The envelope modulation spectra provide the modulation/Power spectrum calculated from energy in extracted amplitude envelopes for original signal and each octave band (125-8000Hz).

The benefit of using this method is that:

a) There is no need to identify vowels and consonant intervals (or to make any linguistic presumptions)

b) Completely automated in MATLAB

c) Can take into account pauses and non-phonetic elements that may occur in the sample (such as dysfluencies in the speech of a PWS).

Keeping in mind the theories put forth to explain the phenomenon of choral reading on stuttering, very few attempts have been made to describe the effects of the fluency inducing condition from the perspective of altered rhythm. This study aims to help understand the ameliorative effects of choral reading on stuttering and attempts to provide an explanation based on the properties of speech rhythm and how they contribute to inducing fluent speech in persons with stuttering.

The questions to be answered in this study include:

1) Does the stuttering frequency reduce across solo reading and choral reading conditions?

2) Does the rhythm, measured using envelope modulation spectra, vary across solo reading and choral reading condition in PWS?

3) Which type of rhythm varies most between the solo reading and choral reading conditions?

CHAPTER 3

METHOD

The objectives of the current study were to:

1. Compare the frequency of stuttering between solo reading and choral reading conditions in Kannada speaking adults who stutter.
2. Compare speech rhythm (using amplitude envelope modulation spectrum) between solo reading and choral reading conditions in Kannada speaking adults who stutter.
3. Identify the AM rates at which maximum changes in rhythm occur between solo reading and choral reading conditions.

Participants

Eighteen Kannada speaking male adults who stutter participated in this study. The age range of the participants was 13 to 33 years (Mean=21.17, SD=5.07). A spontaneous speech sample (monologue/ conversation) of the participants speaking about themselves, along with a reading sample of a standardized Kannada passage was obtained from each participant. Stuttering Severity Instrument for Children and Adults- Third Edition (SSI-3) (Riley, 1994) was administered by the examiner on each participant to determine the severity of stuttering prior to the initiation of the study. The severity of stuttering was determined by calculating frequency of dysfluencies, duration of the blocks, and physical concomitants using the SSI-3, and it was found to range from mild to very severe (mild=3; moderate=7; severe=6; very severe=2). All participants were asked to complete a questionnaire providing their demographic details regarding their native language, age of onset of stuttering, handedness, and family and sibling history. Based on the questionnaire, it was observed that none reported any history of audiological, psychological, neurologic, or other communication disorders. Written consent was taken from all participants who participated in the study. The demographic details of the participants are provided in Table 3.

Table 3

Demographic details of participants.

S.No	Age	Gender	Severity of stuttering	SSI score	Handedness
1	23	Male	Severe	32	Right
2	23	Male	Moderate	25	Right
3	25	Male	Mild	20	Right
4	23	Male	Moderate	27	Right
5	24	Male	Moderate	28	Right
6	33	Male	Severe	33	Right
7	20	Male	Moderate	26	Right
8	28	Male	Severe	33	Right
9	18	Male	Moderate	23	Right
10	17	Male	Mild	21	Right
11	16	Male	Very severe	41	Right
12	16	Male	Very severe	44	Right
13	19	Male	Moderate	25	Right
14	15	Male	Severe	35	Right
15	24	Male	Severe	34	Right
16	19	Male	Mild	17	Right
17	13	Male	Severe	31	Right
18	25	Male	Moderate	26	Right

Sentence stimuli

Forty five Kannada sentences with an average length of 8 words (ranging between 7-11 words) were constructed. Some of the sentences were constructed by translating English sentences from a study by Park and Logan (2015), and by some others by adapting sentences from Kannada books and newspapers. Five native

Kannada speakers rated the sentences based on their familiarity on a 3-point rating scale ranging from not familiar, slightly familiar, and most familiar. Following the familiarity ratings, twenty two sentences rated as most familiar were accepted as the reading material for the solo reading and choral reading tasks.

The twenty two sentences were used for the solo reading experiment. For the external speech signal (recorded voice) used in the choral reading condition, a Kannada speaking adult male with normal fluency participated as the second reader. The twenty two sentences were read aloud into a microphone placed at a distance of 15 cms in a natural speaking manner, which were recorded using the CSL 4000 software-hardware unit. The recording was conducted in a sound attenuated booth. All audio files were then saved in WAV format. The recorded samples were analysed by the examiner using PRAAT software. Extraneous noises were eliminated and any intervals of silence and unnecessary pauses at the beginning and end of the sentence were deleted using PRAAT. These digital recordings were used as the choral speech stimuli in the experiment.

Conditions

Solo reading.

Each participant was asked to first silently read all 22 sentences to familiarise oneself to the stimuli. They were then asked to read the sentences normally without an accompanying signal (recorded voice). The sentences were typed in a 14-point Calibri font and saved as .bmp (image files), and inserted in a Microsoft PowerPoint file for presentation, where each sentence appeared on screen individually after the press of a button controlled by the examiner. The participants were asked to read each sentence without the use of any techniques or changes in their manner of speaking in order to avoid reductions or inhibition of stuttering (Park & Logan, 2015). The sentences were recorded using a digital voice recorder (Sony, ICD-UX533) at a sampling rate of 44100 Hz. The microphone was placed at distance of 15 cms away from the mouth (Boutsen, Brutten & Watts, 2000), and the recorded samples were all saved in WAV format (Park & Logan, 2015).

Choral reading.

For the choral reading condition, each recorded sentence was inserted as an audio file along with a written form of the sentence in a Microsoft PowerPoint file. Each sentence appeared for duration of 6 seconds (the average length of the sentences) with an inter-stimulus interval of 3 seconds. Between each sentence a countdown of 3 seconds appeared in order to prepare the participant for the appearance of the following sentence. The stimuli were presented using a Sony VAIO laptop kept at an approximate distance of 60 cms from the participants' body. The recorded voice (the "second speech signal") was played back to the participant using headphones (Tag- 100 stereo headphones) wherein the sound levels for the headphones were adjusted before the presentation of the stimuli to a level that was considered comfortable to the patient, and was still audible (Kiefte & Armson, 2007).

The participant was asked to begin reading the sentence as naturally as possible without the use of any fluency inducing techniques that would bring about a change in the manner of speaking or result in any inhibition of stuttering. They were first given a trial in order to prepare them for the task. Participants were asked to read along with the recorded voice when they began to hear the reader and were instructed to keep in pace with the reader. They were asked to attempt to read the same words at the same time and to read every word if possible (Ingham &Packman, 1979). In the event of being unable to keep up with the reader, the participant was asked to skip to the part being heard and continue reading. All the reading samples were recorded using a digital voice recorder (Sony, ICD-UX533), placed approximately 15 cm away from the mouth of the participant (Boutsen, Brutten, & Watts, 2000). The recordings were sampled at a rate of 44100 Hz and saved as WAV files.

Data Analysis

Prior to data analysis all the recorded samples with a sampling frequency of 44100 Hz, were converted to mono files from stereo files using Audacity.Ink application, and saved in WAV format. The speech samples were then analysed in the following ways:

a) Frequency of stuttering

The examiner listened to the recorded sentences obtained from the participants from both solo reading and choral reading conditions, and orthographically transcribed them. After that, each syllable was analysed for the presence of dysfluencies. Syllable repetitions, phrase repetitions, blocks, and prolongations were marked according to the classification by Conture (1990). The number of syllables stuttered per sentence was tallied and averaged across 22 sentences to derive the total percent of syllables stuttered (%SS) in the solo reading and choral reading conditions.

Intra-judge and inter-judge reliability

Intra-judge and inter-judge reliability measures were obtained for 20 percent of the recorded samples. The examiner reanalysed 10% of the samples and identified dysfluencies for intra-judge reliability. For the inter-judge reliability another Speech

Language Pathologist, with a minimum experience of five years in dysfluency analysis, analysed 10 % of the samples Cronbach's alpha value was calculated for both intra and inter-judge reliability. The Cronbach's alpha value was found to be 0.99 for intra-judge reliability and 0.98 for inter-judge reliability indicating good consistency as the values are closer to 1.

b) Analysis of rhythm

The analysis of rhythm was carried out by an custom routines (Gnanateja, 2016) which was a modified form of the approach by Leong et al. (2013). The sentences were band pass filtered across five frequency bands using 100th order Finite Impulse Response filters; (1) 100-300Hz; (2) 300-700Hz; (3) 700-1750 Hz; (4) 1750-3950 Hz and (5) 3900-7250 Hz. The amplitude envelope was then extracted by the taking the real components of the Hilbert transform of the signal. The Hilbert envelopes of all the bands were then summed to obtain the broadband envelope of the sentences. The start and end of the envelope were tapered with a 10 ms Hanning window to remove any abrupt amplitude fluctuations at the edges. The envelope was then demeaned and zero padded to obtain a signal length of 20 seconds. The zero padding served as a prerequisite to improve the spectral resolution of the subsequent fast Fourier transform to 0.05 Hz. It was ensured that individual variability in the intensity of the sentences did not affect the peak modulation amplitude. This was done by normalizing the amplitude of the entire modulation spectra by the mean of the entire modulation spectra. The 22 modulation spectra were then averaged, and the peak modulation frequency and amplitude were obtained in three modulation frequency bands (0.9-2.5; 2.5-12; & 12-40).

Two variables were computed for each modulation frequency band, namely the Peak frequency and the Peak amplitude, resulting in six variables (3 rates and 2 variables). The Peak frequency was considered as the frequency of the peak with the largest amplitude and Peak amplitude was the amplitude of the peak normalised to the total energy of the spectrum. The peak frequency and peak amplitude values were extracted into an excel file. Figure 2 shows the envelope amplitude spectrum obtained for one sentence from subject 1 in solo reading condition. On the X-axis is the amplitude envelope amplitude spectrum (dB) and on the Y axis is the amplitude modulation frequency (Hz) ranging from 0 to 50 Hz. Figure 3 shows the envelope amplitude spectrum for one sentence in choral reading condition of subject 1.

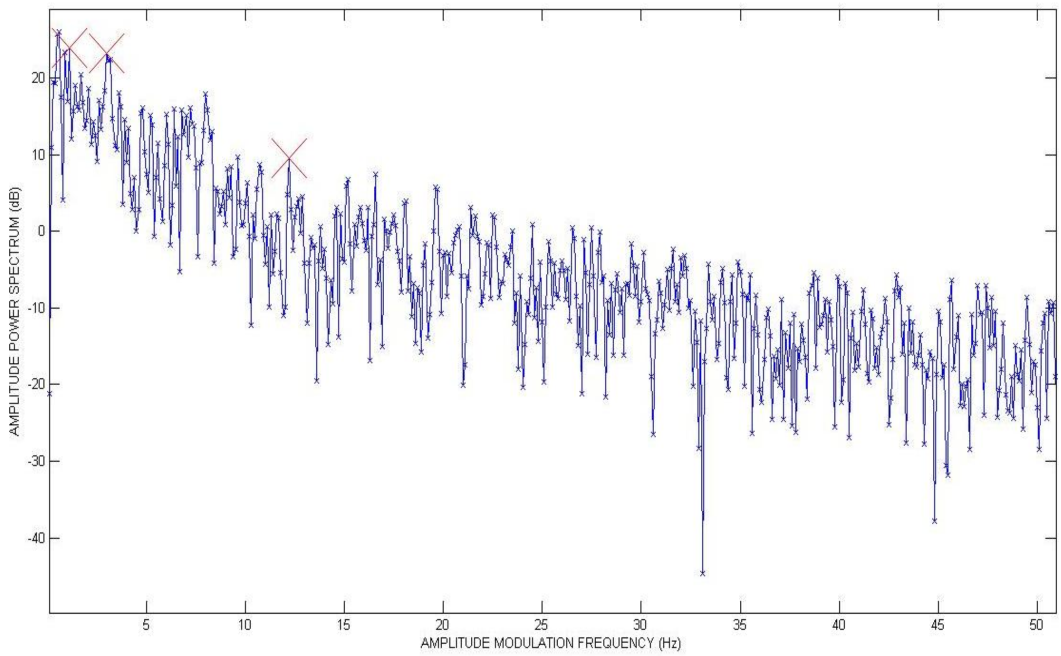

Figure 2: The envelope power spectrum (dB) across three Amplitude modulation rates (Hz) in solo reading condition of Participant 1.

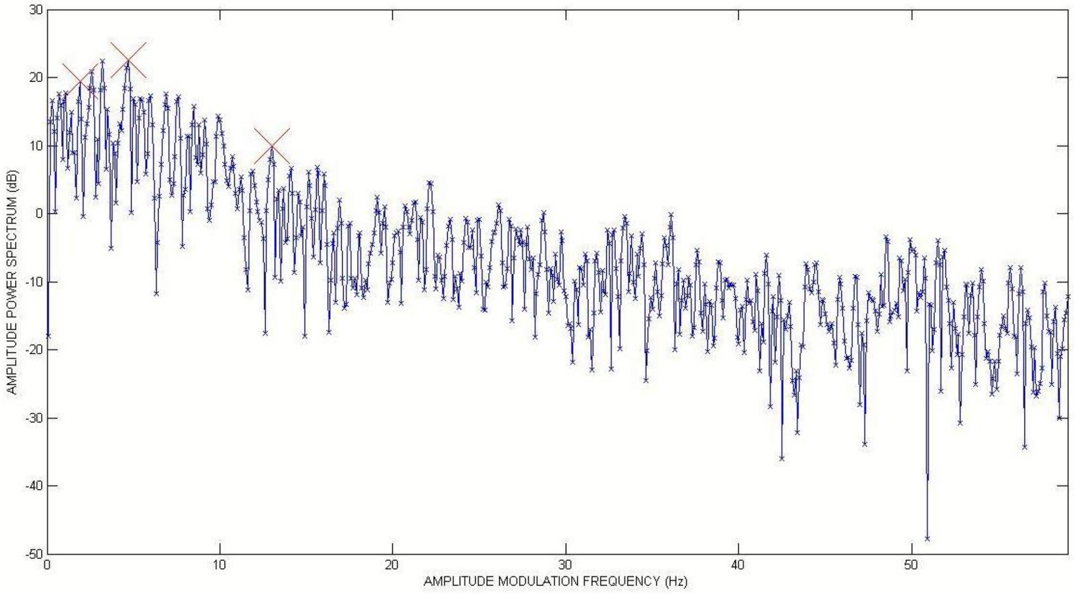

Figure 3. The envelope power spectrum (dB) across three Amplitude modulation rates (Hz) in choral reading condition of Participant 1.

Statistical analysis

All the statistical analysis was done using IBM SPSS statistics (20.0 version). To compare the change in stuttering frequency between two conditions, paired sample t test was used. For the rhythm analysis the dependent measures were considered are peak frequency and peak amplitude across the three rates. A repeated measure ANOVA was performed on the two variables and three rates to study main effects of rate, condition, and interaction between rate and condition. The statistical p value was stated as $p= 0.05$ and effect sizes were reported as partial eta square for ANOVA. For pair-wise comparison of the rates between the solo reading and choral reading conditions, paired t test was done.

CHAPTER 4

RESULTS

The objectives of the current study were to:

1. Compare the frequency of stuttering between solo reading and choral reading conditions in Kannada speaking adults who stutter.
2. Compare speech rhythm (using amplitude envelope modulation spectrum) between solo reading and choral reading conditions in Kannada speaking adults who stutter.
3. Identify the AM rates at which maximum changes in rhythm occur between solo reading and choral reading conditions.

The results of the current study are presented as follows:

1. Comparison of frequency of stuttering (%SS) between solo reading and choral reading conditions

The mean values of percentage of syllables stuttered (%SS) were obtained for each individual for two reading conditions. Paired sampled t test was done to compare % SS between two reading conditions. A significant decrease [t (17) =5.65; p=0.00] in mean percentage of syllables stuttered (%SS) was observed from the solo reading (9.33%) to the choral reading (3.38%) condition. Figure 2 shows comparison of mean percentage of syllables stuttered (%SS) between solo reading and choral reading conditions. In the figure, %SS is on the Y-axis and the solo reading and choral reading conditions on the X-axis. Error bars indicate standard deviation values.

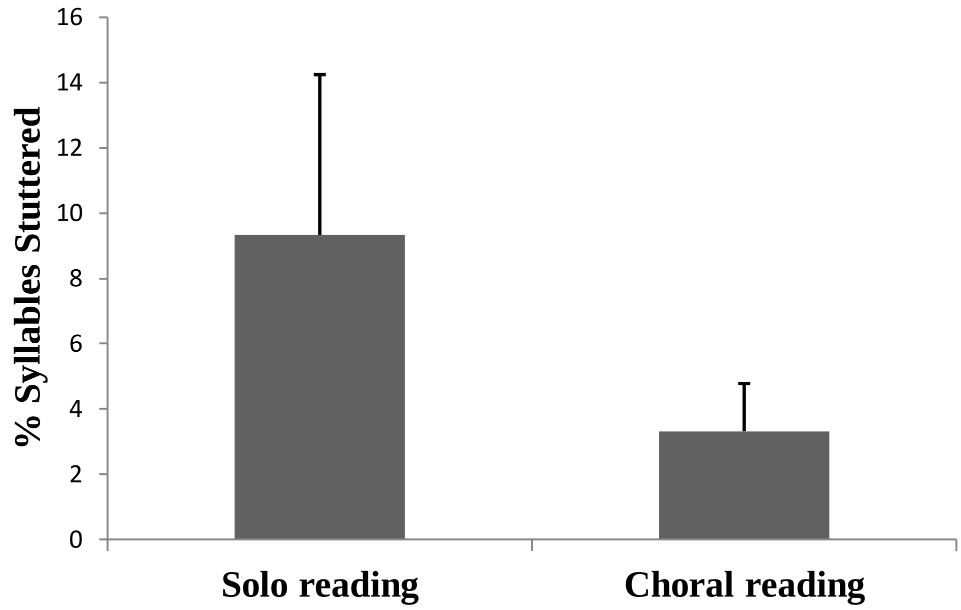

Figure 2. Comparison of %SS between solo reading and choral reading conditions.

2. Comparison of speech rhythm (using amplitude envelope modulation spectrum) between solo reading and choral reading conditions

From the analysis of amplitude envelope modulation spectrum two variables were obtained, 1) Peak frequency; 2) Peak amplitude. The results of the statistical analysis are as follows:

Peak frequency

Mean peak frequency were higher in choral reading condition compared to solo reading condition in two amplitude modulation rates (rate 1 and rate 2 modulation). However, in rate 3 modulation, the mean peak frequency values were comparable between two reading conditions. Repeated measures ANOVA was done to compare mean peak frequency values between two reading conditions. The results revealed statistically significant main effect of rate ($p < 0.05$), condition ($p <0.05$), and interaction between rate and condition ($p < 0.05$). Table 4 shows results of the mean and standard deviation values of the Peak Frequency for three amplitude modulation rates for solo reading and choral reading conditions. Table 5 shows the results of repeated measures ANOVA.

To identify the rate at which greatest difference in Peak frequency values occurred between the two reading conditions, paired t test was done. The results showed significant difference ($p < 0.05$) between solo and choral reading in Rate 1 (Stress rate) and Rate 2 (Syllable rate). However no significant difference ($p > 0.05$) was observed in Rate 3 (Phoneme rate) between both conditions. Table 6 shows results of paired t test for Peak Frequency.

Table 4

Mean and Standard deviation values of Peak frequency at three amplitude modulation rates for solo reading and choral reading conditions.

Rate	**Solo reading**		**Choral reading**	
	Mean	SD	Mean	SD
Rate 1	1.33	0.13	1.52	0.11
Rate 2	3.69	0.41	4.23	0.44
Rate 3	13.65	0.48	13.64	0.45

Table 5

Repeated measures ANOVA results for Peak frequency between conditions.

Source	**df,F**	**Sig.**	**η_p^2**	**Observed Power**
Rate	(1.92,65.46),11440.75	0.000*	0.997	1
Condition	(1,34),9.78	0.004*	0.223	0.86
Rate vs condition	(1.92,65.46),5.29	0.008*	0.134	0.808

Table 6

Results of paired t test for peak frequency.

Rate	df	t	p
Rate 1	34	-4.58	0.000*
Rate 2	34	-3.78	0.001*
Rate 3	34	0.03	0.972

Peak amplitude

Mean peak amplitude values were higher in solo reading condition compared to choral reading condition in two modulation rates (rate 1 and rate 3). However, in rate 2 modulation, the mean peak amplitude values were comparable between two reading conditions. Repeated measures ANOVA was done to compare mean peak amplitude values between solo reading and choral reading conditions across three rates. The results indicated a statistically significant ($p < 0.05$) main effect of rate, condition, and interaction between condition and rate. Table 7 shows results of the mean and standard deviation values of the Peak amplitude for three amplitude modulation rates for solo reading and choral reading conditions. Table 8 shows the results of repeated measures ANOVA.

To identify the rate at which greatest difference in Peak amplitude values occurred between the two reading conditions, paired t test was done. The results showed significant difference ($p < 0.05$) between solo and choral reading in Rate 1 (Stress rate) and Rate 3 (Phoneme rate). However no significant difference ($p > 0.05$)

was observed in Rate 2 (Syllable rate) between both conditions. Table 9 shows results of paired t test for peak amplitude.

Table 7

Mean and Standard deviation of Peak amplitude at three amplitude modulation rates between solo reading and choral reading.

Rate	**Solo reading**		**Choral reading**	
	Mean	SD	Mean	SD
Rate 1	25.21	1.83	22.69	1.38
Rate 2	21.45	1.11	20.99	0.65
Rate 3	8.33	1.05	9.46	0.82

Table 8

Repeated measures ANOVA results for Peak amplitude between conditions.

Source	**df,F**	**p**	**Ƞp 2**	**Power**
Rate	(1.83,62.21)=1345.44	0.000*	0.975	1
Condition	(1,34)=10.98	0.000*	0.244	0.89
Rate vs condition	(1.83,62.21)=17.59	0.002*	0.341	1

Table 9

Results of paired t test for peak amplitude.

	df	**T**	**Sig**
Rate 1	34	4.65	0.000*
Rate 2	34	1.52	0.135
Rate 3	34	-3.6	0.001*

CHAPTER 5

DISCUSSION

The objectives of this study were a) to compare the stuttering frequency between solo reading and choral reading conditions in Kannada speaking adults who stutter, b) to compare speech rhythm between the two conditions using a novel, automatized approach to rhythm analysis i.e., envelope modulation spectra, and 3) to identify the modulation rate at which maximum variation in the rhythm occurred. Results revealed several points of interest. First, the comparison of percentage of syllables stuttered (%SS) was done between solo reading and choral reading conditions. Twenty two sentences were recorded from each participant in both the reading conditions. From these recorded sentences, the analysis of %SS was done. There was significant reduction in the %SS in choral reading condition when compared to solo reading condition. Current results are in consonance with previous studies (Andrews et al., 1983; Freeman & Armson, 1998; Guntupalli, Kalinowski, Saltuklaroglu, & Nanjundeswaran, 2005; Howell & Powell, 1987; Ingham & Carroll, 1977; Ingham & Packman, 1979; Kiefte & Armson, 2008; Rami et al., 2005; Wingate, 1976). Choral speech is a condition where one or more individuals read matching text aloud, together. The effect it has on stuttering has been highly consistent across groups of PWS (Barber, 1939; Cherry & Sayers, 1956; Johnson & Rosen, 1937; Pattie & Knight, 1944) and in individual studies as well (Andrews, Howie, Dozsa, & Guitar, 1982; Freeman & Armson, 1998). Its effects have also been described as dramatic and as observed the reduction in stuttering occurs immediately. Choral speech has been commonly observed in natural situations as well, such as reading together at church gatherings, classrooms etc and has been found to have the same effect on PWS (Bloodstein, 1950).

Second, the comparison of speech rhythm was done between solo reading and choral reading using amplitude envelope modulation spectrum (EMS). The envelope modulation spectrum is the depiction of the slow varying amplitude changes that occur within a signal. It is an automated method that depicts the distribution of energy in the amplitude variations happening across frequencies. Previously, studies have used this method to quantify rhythm alterations occurring in speech (Drullman, Festen, & Plomp, 1994; Greenberg, Arai, & Silipo, 1998). The envelope power spectrum provides the energy distribution occurring throughout a given signal, and is indicative of the minor changes occurring in the amplitude of the signal. In the digital signal processing of speech, every signal behaves as a carrier frequency of its own, and carries multiple amplitude modulations that correspond to different rates. For each sentence, from the extracted EMS, two dependent variables, peak frequency and peak amplitude were obtained across three modulation rates (0.9-2.5 Hz: Stress rate; 2.5- 12 Hz: Syllable rate; 12-40Hz: Phoneme rate). The three AM rates are able to specify the strong-weak rhythm pattern which is commonly found in speech (Leong, 2012). For instance, the 0.9-2.5 Hz corresponding to the stress rate is indicative of imposing a regular stress rate on the utterances thus showing the nature of being "in-sync" with external stimuli (Leong, 2013).

The current results of the rhythm analysis revealed a significant difference in both peak frequency and peak amplitude across all three rates. However, the pair-wise comparison showed that the peak frequency and peak amplitude varied most markedly in the modulation rate corresponding to Stress rate (0.9-2.5 Hz) while the other two rates were variable among the two measures. Current results suggest that PWS may have deficits with their ability to generate regular rhythmic speech gestures at the sentence level. Further, it may be speculated that during choral reading, this deficit in the generation of appropriate rhythmic gestures may possibly getting altered with respect to stressed and unstressed syllables, contributing to the inducement of fluency. The external reader can be assumed to behave as a pace setter for the PWS. Thus, is it possible that the synchronisation of speech between the PWS and the normal reader enhances fluency in turn altering the disrupted rhythm. Going by the explanations provided by Leong and Goswami (2013) the synchronisation of the solo reading and choral reading of the participants being greater at the first rate i.e.,; the stress rate possibly indicates that PWS pace/ match with the external speaker to produce a more “strong-weak” rhythm pattern during choral reading which results in the alleviation of stuttering. As per our knowledge, there are no other published studies on speech rhythm in adults who stutter, and ours is the first study to find evidence for rhythm production changes during choral reading in AWS.

The use of EMS in speech rhythm analysis is novel approach when compared to the conventional methods of analysing rhythm such as the use of rhythm metrics and pair-wise variability indices (PVI). This approach is more appropriate in the analysis of rhythm in disordered speech. This is so because EMS is completely automated unlike the use of manually implemented methods (e.g., PVI), and it holds wide applications in speech and language disorders as it is capable of capturing subtle changes in speech related to treatment. The most important aspect is that EMS measures can handle the non-linguistic parts of the signals like noises, pauses etc (Liss, LeGendre & Lotto, 2010) which are the characteristic features of stuttering dysfluencies. The use of EMS in the field of stuttering is not known of and the analysis of rhythm in PWS is a relatively less studied area. However, it has been used by various authors on other disordered population including dysarthria, dyslexia, and normal infant directed and adult directed speech. Thus, the use of EMS to study disordered speech like stuttering is warranted.

The current findings also support the "disrupted rhythm hypothesis" proposed by Howell, Powell and Khan (1983). This hypothesis specifies that altering voice output influences the speech motor control because it results in a secondary or external rhythmic signal. This secondary signal alters the timing control in individuals with stuttering. Supporting the proposition of disrupted rhythm hypothesis, the current study highlights the use of secondary signals via choral reading and how the use of the second speech signal helps alter the speech timing aspects by altering the rhythm in terms of marked changes in EMS.

The deficit in generation of rhythmic patterns in PWS may be because of deficiencies in the neural resources which are involved in the internal timing of syllables. Evidences from neuroimaging studies suggest that areas which have been found to play important role in rhythmic movement timing are supplementary motor area (SMA), left premotor cortex (PMC), and insula (Grahn & McAuley, 2009; Halsband, Ito, Tanji, & Freund, 1993). Evidences from structural and functional imaging studies also suggest that there is aberrant activity in the above mentioned areas in individuals who stutter (Brown, Ingham, Ingham, Laird, & Fox, 2005; Chang, Horwitz, Ostuni, Reynolds, & Ludlow, 2011; Fox, Ingham, Ingham, Zamarripa, Xiong, & Lancaster, 2000). Given that stuttering individuals in the present study had changes in rhythm between two reading conditions, and previous research suggests role of these areas for rhythmic speech timing, it is possible that abnormal motor circuits of these areas might be responsible abnormal rhythm and choral reading may be changing the activity of these areas.

Evidence is available for this proposition from the past research. Wu, et al. (1995), investigated the neural substrates of choral reading PWS. The persons with severe developmental stuttering were given two tasks, one of reading alone (solo) and the other of reading in unison (choral). They were also compared with control subjects. Significant decrease in the regional glucose metabolism in Broca's area, Wernicke's area, and frontal lobe were noted in the solo reading condition as compared to the choral reading condition which was also observed to be free of stuttering events. Lower levels of activity were observed in the left caudate nucleus (in the basal ganglia) during solo reading and choral reading conditions when compared to the normal subjects. Falk, Muller, and Bella (2015) in their study, highlighted that the stuttering symptoms are eradicated or reduced when the speech production of PWS is paced by an externally provided synchronous tone sequence which is either in pace with another speaker or with a metronome or synced with the speech of another person like in choral reading and shadowed speech. The fluency induced by externally provided rhythm cues has been observed to result in a normalization of hyper- and hypo-activation of neural circuitry that controls temporal processing and initiation of movements namely, the basal ganglia, SMA, and the cerebellum (Toyomura, Fujii, & Kuriki, 2011). It is also possible that a cerebellar mechanisms which focus on multimodal sensory-motor integration is accountable for the voluntary control of timing aspects which influences a number of rhythmic events in PWS.

CHAPTER 6

Summary and Conclusion

In conclusion, the current study aimed at providing an explanation for the induced fluency in PWS during choral reading. Firstly, it was observed that choral reading is successful in marked reductions of stuttering frequency. Secondly, the use of EMS to quantify speech rhythm in solo and choral reading revealed significant changes in the amplitude envelope across three modulation rates, indicating a change in the rhythm type during the choral reading condition when compared to solo reading condition. A significant change was noted in the stress-rate corresponding to the stress-type rhythm. It may be possible that the PWS attempt to correct their disrupted rhythm mechanism by synchronising with the correct mechanism of the external speaker in choral reading. Future research in this field is warranted to study the effects of the "second speech signal" in PWS and whether it can be relied upon during stuttering intervention.

References

Abercrombie, D. (1967). *Elements of general phonetics* (Vol. 203). Edinburgh: Edinburgh University Press.

Adams, M. R., & Ramig, P. (1980). Vocal characteristics of normal speakers and stutterers .

Adams, M. R., & Reis, R. (1971). The influence of the onset of phonation on the frequency of stuttering. *Journal of Speech, Language, and Hearing Research*, *14*(3), 639-644.

Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H., & Merzenich, M. M. (2001). Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proceedings of the National Academy of Sciences*, *98*(23), 13367-13372.

Allen, G. D. (1972). The location of rhythmic stress beats in English: An experimental study I. *Language and Speech*, *15*(1), 72-100.

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

Armson, J., & Kiefte, M. (2008). The effect of SpeechEasy on stuttering frequency, speech rate, and speech naturalness. *Journal of Fluency Disorders*, *33*(2), 120-134.

Balasubramanian, V., Cronin, K. L., & Max, L. (2010). Dysfluency levels during repeated readings, choral readings, and readings with altered auditory feedback in two cases of acquired neurogenic stuttering. *Journal of Neurolinguistics*, *23*(5), 488-500.

Barber, V. (1939). Studies in the Psychology of Stuttering, XVChorus Reading as a Distraction in Stuttering1This study is part of a Ph. D. thesis directed by Dr. Wendell

Johnson at the State University of Iowa as part of a program of studies in the psychology of stuttering. *Journal of Speech Disorders*, *4*(4), 371-383.

Bloodstein, O. (1950). Hypothetical conditions under which stuttering is reduced or absent. *Journal of Speech & Hearing Disorders*.

Bloodstein, O., & Bernstein Ratner, N. (2008). A handbook on stuttering New York: Thomson Delmar Learning.

Bloodstein, O., & Ratner, N. (1995). A handbook on stuttering Singular. *San Diego, CA*.

Boutsen, F. R., Brutten, G. J., & Watts, C. R. (2000). Timing and intensity variability in the metronomic speech of stuttering and nonstuttering speakers. *Journal of Speech, Language, and Hearing Research*, *43*(2), 513-520.

Brown, S., Ingham, R. J., Ingham, J. C., Laird, A. R., & Fox, P. T. (2005). Stuttered and fluent speech production: an ALE meta-analysis of functional neuroimaging studies. *Human brain mapping*, *25*(1), 105-117.

Chang, S. E., Horwitz, B., Ostuni, J., Reynolds, R., & Ludlow, C. L. (2011). Evidence of left inferior frontal–premotor structural and functional connectivity deficits in adults who stutter. *Cerebral Cortex*, *21*(11), 2507-2518.

Cherry, C., & Sayers, B. (1956). Experiments upon the total inhibition of stammering by external control, and some clinical results. *Journal of Psychosomatic Research, 1,* 233-246.

Conture, E. G. (1990b). *Stuttering* (2^{nd} ed). Englewood cliffs, NJ: Prentice-Hall.

Cummins, F. (2003). Rhythmic grouping in word lists: competing roles of syllables, words and stress feet. In *International Conference on Spoken Language Processing* (Vol. 15, p. 28).

Cummins, F., & Port, R. (1998). Rhythmic constraints on stress timing in English. *Journal of Phonetics*, *26*(2), 145-171.

Dayalu, V. N. (2004). *Inhibition of stuttering from second speech signals*(Doctoral dissertation, East Carolina University).

Dellwo, V; (2006) Rhythm and Speech Rate: A Variation Coefficient for delta C. In P. Karnowski & I. Szigeti (ed.), *Language and language-processing* (231-241). http://discovery.ucl.ac.uk/id/eprint/12181

Dellwo, V., & Wagner, P. (2003, August). Relations between language rhythm and speech rate. In *Proceedings of the 15th international congress of phonetics sciences* (pp. 471-474).

Drullman, R., Festen, J. M., & Plomp, R. (1994). Effect of temporal envelope smearing on speech reception. *Journal of the Acoustical Society of America, 95,* 1053–1064.

during choral reading. *Journal of Speech and Hearing Research*, *23*, 457-469.

Eisenson, J., & Wells, C. (1942). A Study of the Influence of Communicative Responsibility in a Choral Speech Situation for Stutterers. *Journal of Speech Disorders, 7,* 259-262.

Falk, S., Müller, T., & Dalla Bella, S. (2015). Non-verbal sensorimotor timing deficits in children and adolescents who stutter. *Frontiers in psychology*, *6*.

Falk, S., Müller, T., & Dalla Bella, S. (2015). Non-verbal sensorimotor timing deficits in children and adolescents who stutter. *Frontiers in psychology*, *6*.

Fox, P. T., Ingham, R. J., Ingham, J. C., Zamarripa, F., Xiong, J. H., & Lancaster, J. L. (2000). Brain correlates of stuttering and syllable production.*Brain*, *123*(10), 1985-2004.

Freeman, K., & Armson, J. (1998). Extent and stability of stuttering reduction during choral reading... winner of the 1997 Isabel Richard Award for most outstanding student paper. *Journal of Speech-Language Pathology & Audiology*.

Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *The Journal of the Acoustical Society of America*, *27*(4), 765-768.

Fry, D. B. (1958). Experiments in the perception of stress. *Language and speech*, *1*(2), 126.

Ghitza, O., & Greenberg, S. (2009). On the possible role of brain rhythms in speech perception: intelligibility of time compressed speech with periodic and aperiodic insertions of silence. *Phonetica., 66*, 113–126.

Glover, H., Kalinowski, J., Rastatter, M., & Stuart, A. (1996). Effect of instruction to sing on stuttering frequency at normal and fast rates.*Perceptual and motor skills*, *83*(2), 511-522.

Gnanateja, G. N. (2016). Rhythm analysis based on Modulation Spectrum. File ID 57199, available at http://in.mathworks.com/matlabcentral/fileexchange/57199-rhythm-analysis-based-on-modulation-spectrum. Covered by BSD open Source license.

Goswami, U., & Leong, V. (2014). Impaired extraction of speech rhythm from temporal modulation patterns in speech in developmental dyslexia. *Frontiers in Human Neuroscience, 8,* 96. doi:10.3389/fnhum.2014.00096

Grabe, E. & Low, E. L. (2002). Durational Variability in Speech and the Rhythm Class Hypothesis. *Papers in Laboratory Phonology, 7*, Mouton.

Grahn, J. A., & McAuley, J. D. (2009). Neural bases of individual differences in beat perception. *NeuroImage*, *47*(4), 1894-1903.

Greenberg, S. (1999). Speaking in shorthand–A syllable-centric perspective for understanding pronunciation variation. *Speech Communication*, *29*(2), 159-176.

Greenberg, S., Arai, T., & Grant, K. (2006). The role of temporal dynamics in understanding spoken language. In P. Divenyi, K. Vicsi, & G. Meyer (Ed.), *Dynamics of speech production and perception* (171–193).

Greenberg, S., Arai, T., & Silipo, R. (1998). Speech intelligibility derived from exceedingly sparse spectral information. *Proceedings of the International Conference on Spoken Language Processing*, 74-77.

Greenberg, S., Carvey, H., Hitchcock, L., & Chang, S. (2003). Temporal properties of spontaneous speech – a syllable-centric perspective. *Journal of Phonetics*, *31*, 465–485.

Guntupalli, V. K., Kalinowski, J., Saltuklaroglu, T., & Nanjundeswaran, C. (2005). The effects of temporal modification of second speech signals on stuttering inhibition at two speech rates in adults. *Neuroscience letters*,*385*(1), 7-12.

Halsband, U., Ito, N., Tanji, J., & Freund, H. J. (1993). The role of premotor cortex and the supplementary motor area in the temporal control of movement in man. *Brain*, *116*(1), 243-266.

Hargrave, S., Kalinowski, J., Stuart, A., Armson, J., & Jones, K. (1994). Effect of frequency-altered feedback on stuttering frequency at normal and fast speech rates. *Journal of Speech, Language, and Hearing Research*,*37*(6), 1313-1319.

Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron, 69*(3), 407-422.

Hirst, D. (2009). The rhythm of text and the rhythm of utterances: from metrics to models. In *INTERSPEECH* (pp. 1519-1522).

Houtgast, T., & Steeneken, J. M. (1985). A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. *Journal of the Acoustical Society of America, 77*(3), 1069–1077.

Howell, P., & Powell, D. J. (1987). Delayed auditory feedback with delayed sounds varying in duration. *Perception & Psychophysics*, *42*(2), 166-172.

Howell, P., Powell, D. J., & Khan, I. (1983). Amplitude contour of the delayed signal and interference in delayed auditory feedback tasks. *Journal of Experimental Psychology: Human Perception and Performance, 9*(5), 772.

Hudock, D. J. (2012). Inhibition of stuttering from second speech signals: An evaluation of temporal and hierarchical aspects. Doctoral Dissertation URI: http://hdl.handle.net/10342/3895

Ingham, R. J., & Carroll, P. J. (1977). Listener judgment of differences in stutterers' nonstuttered speech during chorus-and nonchorus-reading conditions. *Journal of Speech, Language, and Hearing Research*, *20*(2), 293-302.

Ingham, R. J., & Packman, A. (1979). A further evaluation of the speech of Stutterers during chorus- and Nonchorus-reading conditions. *Journal of Speech and Hearing Research, 22,* 784-793.

Johnson, W., & Rosen, L. (l937). Studies in the psychology of stuttering: Effect of certain changes in speech pattern upon frequency of stuttering. *Journal of Speech Disorders, 2,* 105-109.

Kalinowski, J. & Saltuklaroglu, T. (2003). Choral speech: the amelioration of stuttering via imitation and the mirror neuronal system. *Neuroscience Biobehavioural Reviews. 27*(4), 339-347.

Kalinowski, J. S., & Saltuklaroglu, T. (2005). *Stuttering*. Plural Publishing.

Kalinowski, J., Dayalu, V. N., Stuart, A., Rastatter, M. P., & Rami, M. K. (2000). Stutter-free and stutter-filled speech signals and their role in stuttering amelioration for English speaking adults. *Neuroscience Letters*,*293*(2), 115-118.

Kalinowski, J., Guntupalli, V. K., Stuart, A., & Saltuklaroglu, T. (2004). Self-reported efficacy of an ear-level prosthetic device that delivers altered auditory feedback for the management of stuttering. *International Journal of Rehabilitation Research*, *27*(2), 167-170.

Kalinowski, J., Stuart, A., Rastatter, M. P., Snyder, G., & Dayalu, V. (2000). Inducement of fluent speech in persons who stutter via visual choral speech.*Neuroscience letters*, *281*(2), 198-200.

Kiefte, M., & Armson, J. (2007). Dissecting choral speech: Properties of the Accompanist critical to stuttering reduction. *Journal of Communication Disorders, 41,* 33–48.

Kochanski, G., Grabe, E., Coleman, J., & Rosner, B. (2005). Loudness predicts prominence: Fundamental frequency lends little. *The Journal of the Acoustical Society of America*, *118*(2), 1038-1054.

Leong, V. (2013). Prosodic rhythm in the speech amplitude envelope: Amplitude modulation phase hierarchies (AMPHs) and AMPH models. Doctoral dissertation, University of Cambridge.

Leong, V., & Goswami, U. (2013). Assessment of rhythmic entrainment at multiple timescales in dyslexia: evidence for disruption to syllable timing.*Hearing research*, *308*, 141-161.

Leong, V., Hämäläinen, J., Soltész, F., & Goswami, U. (2011). Rise time perception and detection of syllable stress in adults with developmental dyslexia. *Journal of Memory and Language*, *64*(1), 59-73.

Ling, E. E. L., Grabe, E., & Nolan, F. (2000). Quantitative Characterizations of Speech Rhythm: Syllable-Timing in Singapore English. *Language and speech*, *43*(34), 377-401.

Liss, J. M., Legendre, S., & Lotto, A. J. (2010). Discriminating Dysarthria Type from Envelope Modulation spectra. *Journal of Speech, Language, and Hearing Research, 53*, 1246-1255.

Loukina, A., Kochanski, G., Rosner, B., Keane, E., & Shih, C. (2011). Rhythm measures and dimensions of durational variation in speech a). *The Journal of the Acoustical Society of America*, *129*(5), 3258-3270.

Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron. 54,* 1001-1010.

Morton, J., Marcus, S., & Frankish, C. (1976). Perceptual centers (P-centers). *Psychological Review*, *83*(5), 405.

Natke, U., Grosser, J., & Kalveram, K. T. (2001). Fluency, fundamental frequency, and speech rate under frequency-shifted auditory feedback in stuttering and nonstuttering persons. *Journal of Fluency Disorders*, *26*(3), 227-241.

Park, J., & Logan, K. J. (2015). The role of temporal speech cues in facilitating the fluency of adults who stutter. *Journal of fluency disorders*, *46*, 41-55.

Pattie, F. A., & Knight, B. B. (1944). Why does the speech of stutterers improve in chorus reading?. *The Journal of Abnormal and Social Psychology*, *39*(3), 362.

Pike, K. L. (1945). The intonation of American English.

Ptok, M., Natke, U., & Oertle, H. M. (2006). MEDIZIN-Stottern--Pathogenese und Therapie. *Deutsches Arzteblatt-Arztliche Mitteilungen-Ausgabe A*,*103*(18), 1216-1220.

Rami, M. K., & Diederich, E. (2005). Effect of reading with reversed speech on frequency of stuttering in adults 1, 2. *Perceptual and motor skills*, *100*(2), 387-393.

Rami, M. K., Kalinowski, J., Rastatter, M. P., Holbert, D., & Allen, M. (2005). Choral reading with filtered speech: effect on stuttering. *Perceptual and motor skills*, *100*(2), 421-431.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, *73*(3), 265-292.

Riley, G. D., & Ingham, J. C. (2000). Acoustic duration changes associated with two types of treatment for children who stutter. *Journal of Speech, Language, and Hearing Research*, *43*(4), 965-978.

Riley. (1994). *Stuttering Severity Instrument for Children and Adults* (3rd edition ed.). Austin, Texas: Pro-ed.

Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London B Biological Sciences, 336*, 367–373.

Saltuklaroglu, T., & Kalinowski, J. (2006). The inhibition of stuttering via the presentation of natural speech and sinusoidal speech analogs. *Neuroscience letters*, *404*(1), 196-201.

Schane, S. A. (1979). Rhythm, Accent, and Stress in English Words. *Linguistic Inquiry, 10*(3), 483-502.

Scott, S. K. (1993). *P-centres in speech an acoustic analysis* (Doctoral dissertation, University College).

Shannon, R.V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science, 270*, 303–304.

Silverman, E. M. (2003). My personal experience with stuttering and meditation. In *6th Annual International Stuttering Awareness Day Online Conference*.

Skoyles, J. (1998). Speech phones are a replication code. *Medical Hypotheses*, 50, 167-173.

Steeneken, H. J. M., & *Houtgast*, T. (*1980*). A physical method for measuring speech-transmission quality. *Journal of Acoustical Society of Am*erica, *67*, 318-326.

Stuart, A., Kalinowski, J., Armson, J., Stenstrom, R., & Jones, K. (1996). Fluency effect of frequency alterations of plus/minus one-half and one-quarter octave shifts in auditory feedback of people who stutter. *Journal of Speech, Language, and Hearing Research*, *39*(2), 396-401.

Tilsen, S., & Johnson K. (2008). Low-frequency fourier analysis of speech rhythm. *Journal of Acoustical Society of America, 124,* 34-39.

Toyomura, A., Fujii, T., & Kuriki, S. (2011). Effect of external auditory pacing on the neural activity of stuttering speakers. *Neuroimage*, *57*(4), 1507-1516.

Van Riper, C. (1982). *The nature of stuttering*. Prentice Hall.

Villing, R. (2010). *Hearing the moment: Measures and models of the perceptual centre* (Doctoral dissertation, National University of Ireland Maynooth).

Webster, R. L., & Lubker, B. B. (1968). Interrelationships among fluency producing variables in stuttered speech. *Journal of Speech, Language, and Hearing Research*, *11*(4), 754-766.

Wingate, M. E. (1969). Sound and pattern in "artificial" fluency. *Journal of Speech and Hearing Research*, *12,* 677-686.

Wingate, M. E. (1970). Effect on stuttering of changes in audition. *Journal of Speech, Language, and Hearing Research*, *13*(4), 861-873.

Wingate, M. E. (1976). *Stuttering: Theory and treatment*. Irvington.

Wu, J. C., Maguire, G., Riley, G., Fallon, J., LaCasse, L., Chin, S., & Lottenberg, S. (1995). A positron emission tomography [18F] deoxyglucose study of developmental stuttering. *Neuroreport*, *6*(3), 501-505.