Published in final edited form as:

J Fluency Disord. 2011 June ; 36(2): 93–109. doi:10.1016/j.jfludis.2011.03.003.

Systematic Studies of Modified Vocalization: Speech Production Changes During a Variation of Metronomic Speech in Persons Who Do and Do Not Stutter

Jason H. Davidow^a, Anne K. Bothe^b, and Jun Ye^c

^a Department of Speech-Language-Hearing Sciences, Hofstra University, 100B Davison Hall, Hempstead, NY 11549

^b Department of Communication Sciences and Special Education, University of Georgia, 516 Aderhold Hall, Athens, GA 30602

^c Department of Mathematics and Statistics, South Dakota State University, 207 Harding Hall, Brookings, SD 57007

Abstract

The most common way to induce fluency using rhythm requires persons who stutter to speak one syllable or one word to each beat of a metronome, but stuttering can also be eliminated when the stimulus is of a particular duration (e.g., 1 s). The present study examined stuttering frequency, speech production changes, and speech naturalness during rhythmic speech that alternated 1 s of reading with 1 s of silence. A repeated-measures design was used to compare data obtained during a control reading condition and during rhythmic reading in 10 persons who stutter (PWS) and 10 normally fluent controls. Ratings for speech naturalness were also gathered from naïve listeners. Results showed that mean vowel duration increased significantly, and the percentage of short phonated intervals decreased significantly, for both groups from the control to the experimental condition. Mean phonated interval length increased significantly for the fluent controls. Mean speech naturalness ratings during the experimental condition were approximately 7 on a 1–9 scale (1 = highly natural; 9 = highly unnatural), and these ratings were significantly correlated with vowel duration and phonated intervals for PWS. The findings indicate that PWS may be altering vocal fold vibration duration to obtain fluency during this rhythmic speech style, and that vocal fold vibration duration may have an impact on speech naturalness during rhythmic speech. Future investigations should examine speech production changes and speech naturalness during variations of this rhythmic condition.

Educational Objectives—The reader will be able to: (1) describe changes (from a control reading condition) in speech production variables when alternating between 1 s of reading and 1 s of silence, (2) describe which rhythmic conditions have been found to sound and feel the most natural, (3) describe methodological issues for studies about alterations in speech production variables during fluency-inducing conditions, and (4) describe which fluency-inducing conditions have been shown to involve a reduction in short phonated intervals.

© 2011 Elsevier Inc. All rights reserved.

Corresponding author: Jason H. Davidow, Department of Speech-Language-Hearing Sciences, Hofstra University, 100B Davison Hall, Hempstead, NY, 11549, jason.davidow@hofstra.edu. Tel: (516) 463-4582. Fax: (516) 463-4062.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

Stuttering; Fluency-inducing conditions; Rhythmic speech; Vocalization

1. Introduction

Rhythmic speech, most often operationalized as producing one syllable or one word to each beat of a metronome, has been widely studied with persons who stutter (PWS). It typically results in stuttering frequency levels at or near zero (e.g., Andrews, Howie, Dozsa, & Guitar, 1982; Davidow, Bothe, Andreatta, & Ye, 2009). Several explanations have been put forth for the stuttering reductions during metronome stimulation, such as distraction (Barber, 1940; Johnson & Rosen, 1937), slowed speech (Johnson & Rosen, 1937), normalization of brain regions that may be abnormally activated or deficient (Alm, 2004), integration of the speech mechanism (Johnson & Rosen, 1937), or specifically that the metronome effect helps to properly time motoric patterns involved in respiration, phonation, and articulation (Van Riper, 1973), and Wingate's Modified Vocalization Hypothesis (Wingate, 1969, 1970). In this latter hypothesis, Wingate essentially purported that stuttering may be due to irregularities in the rhythm of one's speech and that melodic and prosodic changes during several fluency-inducing conditions (FICs), including metronome stimulation, stabilize these irregularities through inducing continuity in phonation.

In addition to typical syllable- and word-based metronomic speech, other rhythmic patterns can be used to produce fluency in PWS. Jones and Azrin (1969), for example, studied four male participants (19 to 25 years old) who timed their speech to the duration of a vibrotactile rhythmic stimulus applied to the wrist. Stimulus-on durations included 0.1, 0.5, 1, 2, 3, 5, 10, 20, and 300 s; the time between pulses, or the stimulus-off duration, was 1 s in each condition. Group data demonstrated that percentages of words stuttered were at or near zero when participants spoke in the 0.1-s, 0.5-s, and 1-s conditions. After 1 s, the trend was an increase in stuttering frequency, although one participant did not begin to show an increase until after the 3-s condition. Finn and Ingham (1994) used the same conditions to examine how natural various metronome stimulation conditions felt and sounded to PWS. They also found that stuttering levels were at or near zero up to the 1-s condition and increased thereafter.

The rhythmic procedure used in the Jones and Azrin (1969) and Finn and Ingham (1994) studies has received little attention in the literature. This lack of attention is surprising, because both papers reported large reductions in stuttering, and because stimulus-on durations of 1 s to 5 s have been rated as feeling and sounding more natural than shorter durations (0.1 s and 0.5 s) (Finn & Ingham, 1994; Jones & Azrin, 1969). In fact, when rating how their own speech sounded during a condition that alternated 1 s of speech with 1 s of silence, participants in the Finn and Ingham study gave ratings of approximately 5 (1 = highly natural, 9 = highly unnatural; Martin, Haroldson, & Triden, 1984), compared to ratings of approximately 6.5 during a condition that alternated 0.1 s of speech with 1 s of silence. Additionally, stuttering remained at a near-zero level. The improved naturalness with longer stimulus-on durations, coupled with the substantial stuttering reductions, has obvious treatment implications, which warrants further study of this rhythmic speaking style as a possible treatment agent, as suggested by Ingham, Sato, Finn, and Belknap (2001).

A somewhat similar pattern to the ones described in the Jones and Azrin (1969) and Finn and Ingham (1994) studies may have been used in older treatment studies, with investigators trying to "normalize" (likely referring to naturalness) the rhythmic pattern. The most notable example is Brady's (1971) Metronome-Conditioned Speech Retraining (MCSR), in which

speakers went from speaking one syllable or word per beat (the tick of a metronome) at the start of treatment, to speaking “sequences of words to one beat” (Brady, 1971, p. 135) during the later stages of treatment. It is difficult to determine, however, if the sequences of words “normalized” the speech pattern (improved naturalness), as no data are provided in regards to this issue.

The similarity between the speech used in the later stages of MCSR and the rhythmic stimulation with longer stimulus-on durations (hereafter “RSLSD”) used in Jones and Azrin (1969) and Finn and Ingham (1994) is difficult to decipher. That is, no specific descriptions of the speech are provided and the length of the “sequence of words” during the running speech produced at the end of the treatment program is dependent on the metronome rate, information that is also not provided in MCSR studies (e.g., Brady, 1971; Öst, Götestam, & Melin, 1976). Brady does state that speakers typically start the program with the metronome set between 40 and 80 beats per minute (BPM) and that the rate is gradually increased later. Since any BPM rate over 60 would require the sequence of words to be less than 1 s (the beats would be 1 s apart at 60 BPM), it is unlikely that the speakers using MCSR were speaking for 1 s or longer, as was the case in the Jones and Azrin and Finn and Ingham studies.

To the best of our knowledge, the only long-term investigation of RSLSD was presented in the Jones and Azrin (1969) article. One of their participants reduced the percentage of stuttered words from 27.5% to 3.5% (within the clinic) while wearing a portable apparatus (stimulus-on duration was set at 2 s and stimulus-off duration was set at 1 s). The participant and his friends also stated that “his stuttering was almost absent and that his speech sounded natural” (Jones & Azrin, 1969, p. 228) while he used the apparatus. Unfortunately, no data are provided to support this beyond-clinic claim. The reported results are promising, however, and warrant further study of this speaking style.

Additionally, previous research suggests that continued investigation into the utility of rhythmic stimulation as a treatment agent may be fruitful. Andrews, Guitar, and Howie (1980) found metronome stimulation produced the next largest effect sizes for treatment outcome variables after prolonged speech-type procedures during their meta-analysis of treatment studies. Treatment studies using rhythm also received the best ratings for internal validity. It should be noted, however, that in a more recent review (Bothe, Davidow, Bramlett, & Ingham, 2006) only one metronome stimulation study (Öst et al., 1976) met the methodological criteria for inclusion in the review. A recent study also suggests that rhythmic stimulation may be a viable treatment for preschoolers. Trajkovski et al. (2009) found beyond-clinic stuttering reductions below 1 percent syllables stuttered (%SS) at the end of the establishment phase for three preschoolers using one syllable per beat metronomic speech. In addition, the children’s clinicians and parents reported that speech sounded natural beyond the clinic; however, no data were presented to support this claim.

Regardless of the methodological limitations in previous metronome stimulation studies, the fact remains that speaking in time is a powerful fluency inducer, and brain imaging findings show that it can normalize brain regions associated with fluency (Stager, Jeffries, & Braun, 2003). The difficulty with using it as a long-term treatment agent seems to be in the unnatural-sounding speech it produces, especially when the one syllable per beat style is used (Finn & Ingham, 1994); that is, the metronome effect does not carry over (Brady, 1969; Greenberg, 1970), and deviation from the trained rhythmic pattern likely results in a reduction of the fluency-inducing benefits. Therefore, methods for retaining the imposed rhythm that also improve naturalness should be sought.

1.1. The present study

The present investigation was the first in a series of studies designed to investigate RSLSD, with the ultimate objective of this research line being a long-term treatment study using RSLSD as part of a treatment program. Whether or not RSLSD can be used as the foundation of a treatment program (like prolonged speech) will likely depend on whether or not the pattern can be shaped into natural-sounding speech. The present study is an initial step in resolving that issue, and explored changes in speech production variables during RSLSD to identify speech-motor adjustments that may be necessary for fluency induction and that may alter naturalness during this rhythmic speaking style. Changes in speech production variables have been found from a control condition to several FICs, including syllable- and word-based metronomic speech. Examples include an increase in vowel duration (e.g., Stager et al., 1997), an increase in intraoral pressure rise time (Stager et al., 1997), a decrease in the percentage of short phonated intervals (an estimate of the duration of vocal fold vibration, as measured from the surface of the throat between breaks of 10 ms or more; Davidow et al., 2009), a decrease in peak pressure (Stager et al., 1997), and a decrease in air flow rate (Hutchinson & Navarre, 1977) during metronomic speech in a group of PWS. These changes suggest that PWS may receive fluency benefits from these adjustments, although their necessity for fluency during several FICs has yet to be determined (Davidow, Bothe, Richardson, & Andreatta, 2010).

Alternating between 1 s of speaking and 1 s of silence (hereafter “Reading-1.0 s”) was chosen as the experimental condition due to its resulting in near zero stuttering in both Jones and Azrin (1969) and Finn and Ingham (1994); that is, there were notable increases in stuttering frequency with stimulus-on durations longer than 1 s (2 s, 3 s, and 5 s) in both of those studies. We sought to determine if speakers naturally (without instructions to do so) alter aspects of speech production during this variation of metronomic speech. If changes are found to occur naturally, we could determine their necessity to any improvements in fluency in a subsequent study. These initial investigations are critical first steps, as many authors have commented, for example, on how we do not know the most important parameters associated with improved fluency or naturalness for prolonged speech (e.g., Ingham, Montgomery, & Ulliana, 1983; O’Brian, Onslow, Cream, & Packman, 2003). It seems important to identify these changes before examining the utility of a speaking style as a treatment agent.

In addition, naturalness ratings were gathered in the present study. Speech naturalness is most commonly measured in stuttering research using Martin et al.’s (1984) 1–9 naturalness scale. This rating scale has been subjected to substantial investigation as to its reliability and validity. Several studies have shown satisfactory levels of agreement (defined as within 1 scale point) within and between judges (e.g., Martin & Haroldson, 1992; Martin et al., 1984), although others have reported less compelling results (e.g., Onslow, Adams, & Ingham, 1992). Changes in naturalness coinciding with disruptions in the natural flow of speech (Martin et al., 1984), and changes in speech naturalness as speech production variables such as voice onset time, sentence duration (Metz, Schiavetti, & Sacco, 1990), and vowel duration (Schaeffer & Eichorn, 2001) change have also been found. See Schiavetti and Metz (1997) for a comprehensive review of the reliability and validity of the scale.

To date, listeners’ naturalness ratings for RSLSD, using Martin et al.’s (1984) scale, have not been collected. Jones and Azrin (1969) asked listeners to simply mark “natural” or “unnatural,” and Finn and Ingham (1994) used only self-ratings of naturalness by PWS. The ratings provided by the speakers in the Finn and Ingham (1994) study may not be representative of a true listener’s judgment, as there is evidence that PWS rate naturalness differently than other groups of listeners (see Teshima, Langevin, Hagler, & Kully, 2010, for a review). We also examined if speech production changes are associated with different

naturalness ratings during the experimental condition. Finn and Ingham reported standard deviations of approximately 2 during Reading-1.0 s in their study (mean was approximately 5), suggesting a range of scores across the 1–9 scale. It would be informative to see if changes in speech production variables are associated with changes in listener-judged naturalness during Reading-1.0 s, as was the case with, for example, voice onset time during a picture description task (Metz et al., 1990) and vowel duration during the reading of phrases (Schaeffer & Eichorn, 2001). If speech production changes are necessary (or helpful) for fluency during the experimental condition, we need to determine the impact of the changes on naturalness, in order to have a more complete understanding of the utility of Reading-1.0 s as a treatment agent.

1.1.1 Dependent variables—The specific dependent variables examined were vowel duration, voice onset time, fundamental frequency, intraoral peak pressure, intraoral pressure rise time, maximum air flow, vowel midpoint air flow, and phonated intervals (Davidow et al., 2010). These variables were chosen because they represent a global view of the speech system; have changed during other FICs, including syllable- and word-based rhythmic speech (see 1.1. The present study); and/or can be manipulated. For example, phonated intervals have been specifically manipulated (Ingham, Kilgo et al., 2001), vowel durations were altered when focusing on controlling syllable durations (Mallard & Westbrook, 1985), air flow measures can be manipulated using regulated breathing techniques, and, although not experimentally verified, it is reasonable to hypothesize that pressure values are altered when using light articulatory contacts (Stager, Denman, & Ludlow, 1997). Using variables that can be manipulated is important, because of the ultimate goal of identifying variables that can be altered to improve fluency and/or naturalness.

Although vowel duration and phonated intervals are both measures of voicing, they were both included for several reasons. First, the measurements are quite different. The phonated interval is a measurement across syllable and word boundaries whereas vowel duration is not. Second, phonated intervals can be registered without the inclusion of a vowel (voiced consonants); therefore, several phonated intervals may not include a vowel measurement. Third, increases in mean vowel duration do not always correspond with an increase in mean phonated interval duration (Davidow et al., 2010). Lastly, the MPI system provides the ability to obtain hundreds of phonated interval measurements within minutes resulting in the ability to display a distribution of all phonated intervals, whereas, to the best of the authors' knowledge, this technology is not yet available for measures of vowel duration. This large quantity of phonated interval measurements allows for a more detailed and complete analysis of this dependent variable, resulting in findings that would not be possible if only the mean was examined.

1.1.2. Inclusion of normally fluent speakers—Normally fluent speakers were also included in the present study. Data from normal speakers may provide information regarding the importance of speech-motor adjustments for fluency during Reading-1.0 s. If a speech production change is found only for the group of PWS (accompanied by a reduction in stuttering) and not for the group of control speakers, or a more significant change is found, that would provide further support for an association between the speech change and fluent speech during Reading-1.0 s.

1.2. Purposes of this study

The primary purpose of this study was to investigate speech production changes during a condition that alternates 1 s of speaking with 1 s of silence (Reading-1.0 s). We were also interested in listener-judged speech naturalness using Martin et al.'s (1984) 9-point scale and

the association between naturalness and speech production variables during Reading-1.0 s. Therefore, the specific research questions were:

1. Are there changes to speech production variables from a control condition to Reading-1.0 s for PWS and/or normally fluent controls?
2. How will a group of naïve listeners rate the naturalness of the speech produced during Reading-1.0 s?
3. Which speech production variables are associated with naturalness ratings during Reading-1.0 s?

2. Method

2.1. Participants

Thirteen PWS and 11 normally fluent controls participated initially in this study, but several did not meet task compliance criteria (see 2.5. Task compliance). Thus, data from 10 participants were analyzed for each group. The final group of PWS consisted of 7 men and 3 women (mean age = 34.1 years; range = 18–51 years). The final group of controls consisted of 6 men and 4 women (mean age = 31.30 years; range = 20–63 years). One of the PWS had never received treatment; the remaining nine averaged 16.06 years since last treatment (range = 1.5–51 years). No participant reported any past or present neurological disorder or speech, voice, oral-motor, hearing, or head or neck problems that could affect data collection. Previous experience with metronomic speech varied for the PWS. No control group participant had experienced metronomic speech in a formal setting. All participants scored within normal limits on a working memory subtest from the Wechsler Adult Intelligence Scale-Third Edition (Wechsler, 1997), administered as a quick test of the ability to retain a set of instructions.

2.2. Experimental protocol

The data presented in this paper were gathered as part of a larger project consisting of 11 speaking conditions during two visits (Davidow et al., 2010). The speaking tasks presented here were all completed on the first day. The experimental activities began with reading and signing informed consent forms, completing a metronome speech experience questionnaire, and completing the working memory subtest. Speaking tasks were then initiated. The experimental protocol was the same for participants in both groups.

2.2.1. Speaking tasks—The first two speaking tasks were monologues (four 3-min trials) and oral reading from a high school-level textbook (four 3-min trials). The order of these two tasks was reversed for each subsequent participant. These tasks provided habitual stuttering rates for PWS and speech rates for both groups. In addition, these values were used to ensure that carryover between conditions was minimized (see 2.2.2. Rest periods).

After these initial tasks, the main data-collecting conditions were initiated (see Table 1 for an example of the order of the speaking conditions). The two conditions (control and experimental) presented in the present study were included in the randomization of the 11 speaking conditions in the larger project (Davidow et al., 2010), with the exception that they had to be performed on the first day. This exception was necessary for proper implementation of later conditions in the larger project. Five out of 10 control group participants and 4 out of 10 PWS completed the control condition first. The control and experimental conditions involved reading aloud from a high school-level textbook for two 75-s trials, during which acoustic and phonated interval data were collected, and three 30-s trials during which aerodynamic data were collected. The sets of trials for each condition occurred one after the other and the order of the sets was reversed for each subsequent

participant (half did 75-s trials first and half did 30-s trials first). Carrier phrases (see Appendix) were embedded throughout the text to ensure production of specific target words. The same carrier phrases were used for the control and experimental conditions. The carrier phrases, and text sections within which the carrier phrases were embedded, were randomized across the conditions.

The two speaking tasks analyzed for the present study included (a) control reading, and (b) reading during evenly spaced stimuli of 1.0-s duration. During the control reading condition, the participants read aloud and were instructed to read as they normally would. No speaking style or speech rate was prescribed. The experimental condition, Reading-1.0 s, involved alternating between 1 s of reading aloud and 1 s of silence; that is, the stimulus-on duration was 1 s and the stimulus-off duration was 1 s. No speech rate or speaking style was prescribed beyond the following instructions: match your reading with the tone; that is, “read while the tone is on and do not read while the tone is off.” The participant would start reading when the tone was on, stop reading when it went off, and continue at the spot in the text where they stopped when the tone returned.

Participants practiced the Reading-1.0 s condition until the experimenter (first author) judged their speech to score “1” or “2” on a 7-point (1–7) scale, with “1” defined as *definitely producing this condition correctly* and “7” defined as *definitely not producing this condition correctly*, during two consecutive 30-s practice trials. The experimenter listened for synchronization of the participants’ utterances with the 1.0-s tone. If the participants did not meet this pre-condition task compliance criterion after six trials, the condition was initiated anyway. This occurred for only one participant. His data were excluded because his performance did not meet post-experiment task compliance criteria. Text sections used during the practice trials were different than those used during the control reading and Reading-1.0 s conditions.

2.2.2. Rest periods—Two-min monologues were placed after each speaking condition to prevent carryover effects (Davidow et al., 2009). Before the next condition could begin, PWS were required to reach %SS levels within 90% or greater of their baseline monologue rate, reach 90% or greater of their baseline monologue speech rate, and receive a rating of “2” or lower on a 7-point (1–7) scale (“1” = *sounds identical to habitual monologues*; “7” = *sounds totally different than habitual monologues*). Control group participants were only required to reach the latter two criteria. These data were collected online by the experimenter and verified post-experiment by a graduate student who had recently completed a stuttering measurement training program (Ingham & Ingham, 2004), which uses a perceptual threshold definition (Martin & Haroldson, 1981) to teach the counting of stuttered syllables. Post-experiment data showed that all participants who required rest period data met the criteria prior to beginning the two conditions in the present study, with an average of 1.13 2-min monologues before Reading-1.0 s for the PWS (n = 8) and 1.0 for the controls (n = 9), and an average of 1.22 2-min monologues before control reading for the PWS (n = 9) and 1.0 for the controls (n = 8). The number of participants was less than 10 because those conditions randomized to the first position were not preceded by a rest period.

2.3 Acoustic and Phonated Interval Data

Acoustic and phonated interval data were collected during two 75-s trials in a sound-treated room. The 1-s 250 Hz tone for the Reading-1.0 s condition originated from a laptop computer and was presented through a unilateral insert earphone in the participant’s right ear. The intensity of the tone was established before the first set of trials (acoustic or aerodynamic) and the same intensity was used across the sets of trials. It was specific to each participant, and was set by adjusting the computer volume controls until the lowest

level that could be heard through the earphone while performing the speaking task comfortably was found. This resulted in a mean intensity level of 47.38 dB SPL (range = 46.5 – 50 dB) across the groups.

Acoustic data were collected using the PRAAT version 4.3.27 acoustic analysis program. The acoustic signal was transduced with an AKG-C420 head-mounted condenser microphone placed approximately 6 cm from the left oral angle. The microphone was connected to a Mackie Micro Series 1202 12-Channel Mic/Line Mixer which was then connected to the input of a PC sound card. Input gain was set to 50% of the total operating range. The sampling rate was set at 22 kHz. Participants read for approximately 15 s with no data collection, followed immediately by PRAAT and phonated interval recording functions and commencement of a 75-s trial.

Acoustic dependent variables included vowel duration, voice onset time, and fundamental frequency and were gathered from target words in the carrier phrases. Measures were gathered using the PRAAT waveform display, accompanying spectrogram, and audio-recording functions. Vowel duration was defined as beginning at the onset of periodic variation in the waveform following the voiceless consonant and ending at the offset of periodic vibration. Voice onset time was defined as the time from the onset of the stop consonant burst to the onset of voicing. Vowel duration measures came from the first vowels in each target word (see Appendix), except for *spectators* and *spectacular* (the first and second vowel were measured), and *attack* (only the second vowel was measured). Voice onset time measures were from the initial sound in the target words that begin with /p/, /t/, and /k/, in addition to the /t/ in *attack*. Fundamental frequency measurements were gathered by finding the average within the vowel used to collect vowel duration measures.

As was also the case for aerodynamic data, improperly produced and stuttered target words were discarded. Variations in dialect were allowed. In addition, if any word adjacent to a target word was judged as improperly produced or stuttered, that target word was discarded. The first author and a graduate student who completed a stuttering counting training program (Ingham & Ingham, 2004) watched all video recordings. If either of them judged a target word or an adjacent word to be produced improperly or stuttered, that target word was discarded.

Phonated intervals were collected using the Modifying Phonated Intervals (MPI) software program (Ingham, Moglia, Kilgo, & Felino, 2006). The MPI system (Windows platform) consisted of a single-axis accelerometer (Measurement Specialists, Huntsville, AL; Model ACH-01-04), a signal conditioning amplifier, customized software, and related hardware. The accelerometer was a piezo-electronic transducer with no sensitivity relative to the earth and a frequency response of 2 Hz to 20 kHz. The MPI system uses a Sound Blaster Live card (16-bit) with a frequency response of 10 Hz-44 kHz. The signal from the accelerometer is bandpass filtered at 80–300 Hz (7th order Butterworth). The system sets the digitization rate for the sound card at 12 kHz yet ignores 11 of each 12 samples, for an effective sampling rate of 1 kHz. Input signals were integrated and smoothed over a 10-ms window, after which only alternate values were kept, resulting in an effective sampling rate of 500 Hz for the conditioned signal. Before the start of each MPI session, the system noise floor was established by the user. System noise was typically between 62 and 75 mV of noise maxima. When the intensity of the incoming signal exceeded the noise floor by greater than 10%, the recording of a phonated interval was initiated by the analog/digital system. In addition, the measurement of a phonated interval continues when there is a break in voicing for less than 10 ms. This was done to eliminate head movements or vocal fry from registering a phonated interval (R. J. Ingham, personal communication, February 6, 2011).

The phonated interval fitting protocol was completed as follows. Participants were fitted with the accelerometer that was held by an elastic collar and worn comfortably around the neck such that the accelerometer was paralateral to midline and just inferior to the thyroid prominence. The MPI program was then engaged, sampling gain was adjusted, a background noise check was conducted, and movement artifact checks were completed (e.g., participants were asked to nod; if undesired signals were registered, the accelerometer was repositioned until head movements did not record a phonated interval). Participants were instructed not to cough, clear their throats, laugh, or make other similar movements during data acquisition trials, because these behaviors might register a phonated interval. Reliability and validity (agreement with an acoustic analysis program) data for the MPI system can be found in Davidow et al. (2009).

Participants were instructed to maintain their vocal intensity during the 75-s trials within a 4-dB range that was consistent with their performance during a 1-min control reading trial. Vocal intensity was monitored via the LED display on the Mackie Micro Series 1202 12-Channel Mic/Line Mixer and if necessary, the experimenter signaled the participants to adjust their volume to stay within the 4-dB range.

2.4. Aerodynamic Data

Participants were seated in a room with a custom designed aerodynamics workstation. Once seated for a set of trials, the participants were told to remain still and limit their head movement until the experimenter signaled the end of the set. The experimental condition was paced via a tone emitted from the same laptop computer used for the acoustic and phonated interval data. The participants could choose either ear for the unilateral insert earphone (either ear was allowed, as opposed to during the acoustic trials where the structure of the microphone required using the right ear).

2.4.1. Instrumentation—Intraoral pressures, via a Honeywell® Microswitch pressure transducer (Model 164PC01D37), and air flow measures, via a full-face mask coupled to a pneumotachometer (Hans Rudolph® Model R4719), were collected. A 5-centimeter segment of plastic tubing (Intramedic PE 260) (taped to the face mask) was coupled to the transducer and threaded through the face mask for the pressure measures (see Barlow, Suing, & Andreatta, 1999 for a drawing of this setup). The mask was held tightly against the participant's nose and mouth so that the tube's aperture was orthogonal to the corner of the mouth and behind the lips. Judges reported no difficulty counting syllables and stutters when the mask was in place.

Intraoral pressures and air flow were recorded using a Windows-based custom-designed software package known as AEROWIN (Neuro Logic, Lawrence, KS; for specifics see Barlow et al., 1999). All signals were conditioned and filtered by a bridge amplifier (LP -3 dB @ 50 Hz, Butterworth 4-pole, DC coupled, Biocommunication Electronics Model 225), and routed to a 16-bit deglitched analog-to-digital converter (National Instruments, Inc.). Audio signals were collected via a Sennheiser e815S microphone placed approximately 10–12 inches from the participant's mouth. The microphone distance from the participant's mouth was kept constant across conditions. The air flow and air pressure transducers were calibrated prior to each experimental data collection session. The system was calibrated for air flow using a rotometer set to 400 cubic centimeters per second (cc/s) and for pressure using a u-tube manometer to displace 10 cmH₂O.

2.4.2. Data collection—Two of the six carrier phrases (see Appendix), embedded within the reading text, were produced during each 30-s trial. The AEROWIN program uses a 6-second data acquisition window; therefore, the participants began to read and when they

reached the first carrier phrase, they paused and the face mask was placed against their nose and mouth. The face mask was removed after 6 s of reading, as the participant continued to read. This process was repeated for the second carrier phrase in the trial, and the participant continued to read until the experimenter signaled the end of the 30-s trial.

Participants were instructed to maintain their vocal intensity within a 4-dB range identified via a sound level meter during a 1-min oral reading task before the first aerodynamic trial. If the participant deviated from the vocal intensity range more than 3 times outside the data acquisition window, the entire trial was repeated. This occurred on three occasions. Verbal reminders were provided to the participants during the trials if necessary.

2.4.3. Measurement—The four dependent variables measured using the AEROWIN system were peak intraoral pressure, pressure rise time, maximum air flow, and vowel midpoint air flow. As was the case with the acoustic variables, targets were embedded in the carrier phrases. The first author, using the graphic display of the AEROWIN system, gathered the measurements. Initial identification of each target from the recordings was completed by visual inspection and using the time-locked audio recording. Peak intraoral pressure (displayed in cmH₂O) values were obtained by measuring the maximum pressure value for /p/ and /b/ sounds at the beginning of words in the carrier phrases. The target sounds were surrounded with vowels to allow easy identification of the target in the signal and to set intraoral pressure at a minimum before target production (Hutchinson & Navarre, 1977; Stager et al., 1997; Stager & Ludlow, 1993). Pressure rise time (displayed in ms) was obtained by calculating the duration from a point immediately before the pressure rise for /p/ or /b/, to the point of peak pressure. Maximum air flow (displayed in cc/s) was defined as the highest point on the air flow trace after the release of the /p/ or /b/. Vowel midpoint air flow (displayed in cc/s) was gathered for the vowel following the /p/ and /b/ productions by identifying the center of the vowel's production duration. An example of the aerodynamic signals can be found in Figure 1.

2.5. Task Compliance

Acoustic task compliance measures were gathered after data collection had been completed for the Reading-1.0 s condition. The first author measured the length of time the participant read during stimulus-on durations and the length of silence during stimulus-off durations, during two randomly selected 10-s periods for each acoustic trial. If 10 measurements (5 stimulus-on durations and 5 stimulus-off durations) were not obtained in the selected 10-s period, the judge continued until 10 were gathered. The mean value was calculated for both measurements for each trial. To be included in data analysis, the average length of utterance and the average length of silence were required to be within 300 ms of 1.0 s (0.7 to 1.3 s) for each trial. If one or more of the values for either trial was outside of this range, that participant was not included in the final data analysis.

2.6. Stuttering frequency and speech rate

Percent syllables stuttered and syllables per minute (SPM) data were collected for all trials from video recordings by the first author using the Stuttering Measurement System (SMS; Ingham, Bakker, Ingham, Moglia, & Kilgo, 2005) software, which allows an observer using a computer mouse to count syllables spoken with and without stuttering. Judges used a perceptual definition of stuttering (Martin & Haroldson, 1981) to identify stutters. SPM was defined as the number of syllables produced in a trial divided by the length of the trial. A second measure of speech rate, articulation rate, was collected from the 75-s trials. This measure was necessary because the mandated pausing during the Reading-1.0 s condition produced artificially slow speech rates using the first calculation. Articulation rate was calculated by measuring the amount of time and number of syllables produced between

pauses. The SPM value for this measure was found by multiplying the number of syllables by 60, and dividing that value by the amount of time reading (in seconds) between pauses. Pauses and the amount of time between pauses were identified via the PRAAT waveform display, accompanying spectrogram, and audio-recording functions. Pauses were identified as an absence of phoneme production, which may or may not contain an audible inspiration, in the acoustic waveform greater than 150 ms. The 150-ms marker has been used in previous studies with reading (Hammen & Yorkston, 1996; Maassen & Povel, 1984). To assist in the identification of pauses, judges were told that pauses often accompany an audible inspiration. The judges were an upper-level undergraduate student and a graduate student. Each judge finished a complete data set (both conditions) for half of the participants. For PWS, the articulation rate was gathered using only stutter-free segments of speech, as the inclusion of stuttered segments may impact the articulation rate measure. For this measure, the first author identified which 5-s segments contained a stutter, allowing for measurement of a stutter-free articulation rate using only those 5-s segments not containing a stutter.

2.7. Speech naturalness

Eleven graduate students rated the naturalness of the 75-s reading samples. The group consisted of 2 men and 9 women who did not have previous experience rating speech naturalness. Key elements of their instructions, which were adapted from Martin et al. (1984), included the following:

You will hear a number of 1.25-minute reading samples... Your task is to rate the naturalness of each Make a rating halfway through the sample (reflecting the naturalness of the first half of the sample) and at the end of the sample (reflecting the naturalness of the second half of the sample). If the reading sample sounds highly natural to you, circle the 1 on the scale. If the reading sample sounds highly unnatural, circle the 9 on the scale. If the reading sample sounds somewhere between highly natural and highly unnatural, circle the appropriate number on the scale. Do not hesitate to use the ends of the scale (1 or 9).... "Naturalness" will not be defined Make your rating based on how natural or unnatural the reading sounds to you.

The order of the samples was randomized across judges. Each judge re-rated all samples with at least two weeks between sessions, using a different random order. Only data from those judges who exhibited 80% intrajudge agreement were used in the final data analysis. Agreement was defined as the two ratings being within one scale point (Finn & Ingham, 1994; Martin et al., 1984). Eight of the 11 judges met this criterion. Each speaker's naturalness score was the average of 32 scores (8 judges x 2 ratings for each of 2 samples).

2.8. Data Analysis

A set of Cochran-Mantel-Haenszel (CMH) tests (Wickens, 1989), which is a family of tests aimed at detecting a direct association between two variables after controlling for the third variable (usually called the stratified variable), was conducted. The null hypothesis of the CMH test is that there is no association, or no pattern of association, between the two tested variables across the stratified variable. In the present study's case, the two tested variables were condition (control reading and Reading-1.0 s) and speech variable (mean vowel duration, mean voice onset time, mean duration of phonated intervals, mean fundamental frequency, mean peak pressure, mean pressure time, mean maximum air flow, and mean vowel midpoint air flow), and the stratified variable was participants. The stratified variable was used as a control variable, so that the effect of participants that could influence the relationship between condition and the dependent variables could be controlled. In summary, the CMH tests were used to evaluate the strength of association between the dependent variables and their different conditions after controlling for the stratified variables.

(i.e., different participants in the study). A similar analysis was used in one of our previous FIC investigations comparing the same dependent variables between two conditions (Davidow et al., 2010).

CMH test is a methodology based on the randomization model approach. For the one-sample repeated measurements problem, CMH test has advantages over the traditional repeated measures MANOVA method. The randomization model approach performs well when the data sample sizes (i.e., the numbers of subjects) are too small to warrant the use of other large-sample methods (Davis, 2002). Specifically, a CMH test is more appropriate in situations where the data sample size is small and the number of repeatedly measured dependent variables is relatively large, which was the case in the present study. A CMH test does not require random sampling of subjects from some underlying probabilistic framework and could be used in a relatively assumption-free context (Davis, 2002).

An alpha level of .05 was used for statistical significance for the CMH tests, as well as all statistical tests, due to the preliminary nature of this investigation and the relatively small sample size. For small sample sizes, although procedures such as Bonferroni correction avoid occasional Type 1 errors, they can exacerbate the problem of low power and inflate Type II error to the point of obscuring important and interesting findings (Nakagawa, 2004). Additionally, effect sizes and family-wise error rates ($\text{FWER} = 1 - [1 - .05]^c$, where c = number of comparisons per family) are provided for a complete picture of the experimental and statistical effects. The FWER for the family of CMH tests just discussed was .10. “Families” were categorized by different hypotheses or different groups in statistical inference (Hochberg & Tamhane, 1987; Tothaker, 1991).

Two CMH tests were also conducted for the phonated interval distributions, since the MPI system allows for the collection of a considerable number of phonated intervals, resulting in a more fine-grained analysis of this dependent variable. A similar analysis was conducted between control conditions and FICs in our previous FIC and phonated interval investigation (Davidow et al., 2009). For this analysis, the test of association was between the condition and phonated interval bins. The primary dependent variable was the percentage of phonated intervals that fell into 20 bins, with one 70-ms (30–100 ms) duration bin, 18 50-ms (from 51–1000 ms) duration bins, and one duration bin encompassing all phonated intervals above 1000 ms. As was the case for the CMH tests involving acoustic and aerodynamic variables, the distribution of phonated intervals was compared between control reading and Reading-1.0 s for both groups ($\text{FWER} = .10$).

Analyses of individual dependent variables were also completed. Each participant produced several tokens for each dependent variable, resulting in a mean for each. These means were compared using paired-sample t -tests. For the phonated interval bin analysis, each participant had a percentage score and those scores were subjected to a t -test. Phonated intervals in the range of 30–200 ms (one 70-ms bin and two 50-ms bins) were subjected to a t -test, since previous studies suggest alterations in the number of phonated intervals produced in this range can control stuttering (Gow & Ingham, 1992; Ingham et al., 1983). FWER was .43 for each group’s set of tests. Effect sizes (*Cohen’s d*) were also calculated for each dependent variable. Additional t -tests were run to determine if there was a significant difference in the change from the control to experimental condition between the two groups for the individual independent variables ($\text{FWER} = .43$). Lastly, Pearson correlation coefficients were conducted to examine the association between the change in articulation rate and the change in those variables that were found to be significantly different between conditions ($\text{FWER} = .19$), and between naturalness ratings (using the average rating across both trials) and speech production variables (the eight main dependent

variables and the three bins in the 30–200-ms range), for each group (FWER = .43 for each group's set of tests).

2.9. Reliability

Approximately fifteen percent of the data were rejudged for interjudge and intrajudge reliability. The reliability judge for stuttering frequency, SPM, and 5-s stuttered segments for calculating stutterfree articulation rate was a graduate student who had recently completed a stuttering measurement training program (Ingham & Ingham, 2004). The reliability judges for the acoustic, aerodynamic, and task compliance data were students trained specifically for those tasks by the first author. The reliability judge for the articulation rate calculations was the first author.

3. Results

3.1. Reliability

Reliability data for all dependent variables, with the exception of %SS and articulation rate, are shown in Table 2. Two measures were used to assess reliability: mean difference (between judges or between measurement occasions) and the mean of percent deviation scores (difference scores for each token were divided by the primary judge's [or initial] measurement, and multiplied by 100). The average percent deviation for interjudge data across all dependent variables in Table 2 was 4.14. The average for intrajudge data was 1.83.

Interjudge reliability for stuttering frequency was gathered via 33 randomly selected 3-min reading and monologue trials. Nineteen of the 33 trials differed by 0.5 %SS or less, eight differed between 0.5 and 1.0 %SS, five between 1.0 and 1.5 %SS, and one between 2 and 3 %SS. The average difference for the 10 trials that contained the most stuttering (mean = 10.80 %SS) was 0.91 %SS. Intrajudge reliability was gathered via 25 randomly selected 3-min reading and monologue trials. Seventeen of the 25 trials differed by less than 0.5 %SS, five differed between 0.5 and 1.0 %SS, two between 1.0 and 1.5 %SS, and one between 2.0 and 2.5 %SS. The average difference for the ten trials with the most stuttering (mean = 11.43 %SS) was 0.79 %SS. Interjudge reliability data for the 5-s segments with a stutter show that 93.94% (31/33) of the 5-s segments that were identified by the primary judge as stuttered, were also judged as stuttered by the reliability judge. Two 5-s segments identified as stuttered by the reliability judge were not identified as stuttered by the primary judge.

Reliability data for speech rates using the SMS system were collected by randomly selecting 3-min reading and monologue trials. For interjudge reliability, 10 of the 33 trials selected differed by 5 SPM or less, three differed between 5 and 10 SPM, 10 between 10 and 15 SPM, three between 15 and 20 SPM, and seven between 30 and 40 SPM. For intrajudge data, 11 of 25 randomly selected 3-min reading and monologue trials differed by less than 5 SPM, six differed between 5 and 10 SPM, five between 10 and 15 SPM, and three between 20 and 30 SPM.

Interjudge reliability data for articulation rate were gathered by re-analyzing the acoustic trials for 12 randomly selected trials, consisting of three control reading and three experimental condition trials from the data sets produced by each judge. Of the twelve trials analyzed, three differed by less than 1 SPM, six between 1 and 4 SPM, and three between 6 and 8 SPM. For articulation rate intrajudge reliability, five trials were re-judged by each rater, consisting of two control reading trials and three experimental condition trials. Five trials differed by less than 1 SPM, three between 1 and 2 SPM, and two between 3 and 4 SPM. These data show that the primary judges identification of pauses, the primary measure used to calculate articulation rate, were reliable. Overall, the interjudge and intrajudge

agreement data for all dependent variables were considered satisfactory to support the interpretations presented.

3.2. Task Compliance

Three of 13 PWS and 1 of 11 controls did not meet the task compliance criteria; thus, data are presented for 10 participants in each group. The mean task compliance values for the stimulus-on portion were 1.11 s (range = 1.01–1.29 s) for the controls and 1.09 s (range = 0.84–1.25 s) for the PWS. The stimulus-off means were 0.92 s for both groups (controls = .76–1.01 s; PWS = 0.77–1.17 s).

The aerodynamic data were collected in 6-s epochs in which the carrier phrases were produced. For several reasons, some target words were not produced by all speakers during all aerodynamic trials. The number of words was matched for final analyses so that the same words were included in data analysis for both conditions, for each participant. This resulted in a mean of 11.7 words for the control group and 9.2 words for the group of PWS. These averages are well above those used during the latest peer-reviewed FIC and vocalization study that analyzed aerodynamic data (Stager et al., 1997). These adjustments were not necessary for the acoustic data, because all target words were attempted during the 75-s trials.

3.3. Stuttering frequency, speech rate, and speech naturalness

Table 3 shows the SPM, articulation rate, %SS, and speech naturalness data. The rhythmic condition produced a large drop in stuttering from all normal speaking conditions, resulting in 0.28 %SS across the group of PWS. Individual data show that 8 out of 10 PWS did not stutter during the experimental condition, one averaged 2.5 %SS, and one produced two stutters during one of the acoustic trials. The participant averaging 2.5 %SS produced 16.24 %SS during control reading, so the experimental condition still resulted in a substantial drop in stuttering. Table 3 also reveals that, for both groups, the experimental condition was performed using a slower rate of speech than the control condition. Naturalness ratings were essentially identical across groups during Reading-1.0 s.

3.4. Multiple Variable Analysis

CMH tests were performed comparing the distributions of the eight dependent variables between control reading and Reading-1.0 s for both groups. The tests for the controls (CMH test value = 127.17; $p < .0001$) and PWS (CMH test value = 42.23; $p < .0001$) were significant, showing that the distribution of the variables was different during Reading-1.0 s than during the control condition.

3.5. Individual Variable Analysis

Paired sample *t*-tests were conducted to examine whether or not there was a change in the individual dependent variables from the control condition to Reading-1.0 s. Table 4 displays the mean values and effect sizes for each dependent variable. The *t*-tests for vowel duration were significant at the 0.05 significance level for the controls ($t = -2.54$) and PWS ($t = -2.35$), with longer mean vowel durations during Reading-1.0 s for both groups. Effect sizes were above 0.7 (for Cohen's *d* 0.2, 0.5, and 0.8 are considered small, medium, and large effect sizes, respectively), and individual data showed that 9 of the 10 participants in the control group, and 7 out of 10 of the PWS, had longer mean vowel durations during Reading-1.0 s. Lastly, the control group had a significantly longer mean phonated interval length (mean length of all phonated intervals) during Reading-1.0 s than during control reading ($t = -3.15$, $p = .01$, effect size = .80). Eight out of ten participants in that group had longer mean phonated interval lengths during Reading-1.0 s. The aerodynamic variables,

voice onset time, and fundamental frequency did not change significantly between conditions. The similarity between peak pressure and air flow measures, which are known to be positively correlated with vocal intensity, suggests that the participants used a relatively constant vocal intensity across conditions. All *t*-tests examining the difference in the change from the control condition to Reading-1.0 s between the two groups were not significant.

3.6. Phonated interval distribution analysis

The CMH tests examining differences in the distributions of phonated intervals between control reading and Reading-1.0 s were significant for the controls (CMH test value = 33.54; $p = .02$) and the PWS (CMH test value = 49.37; $p = .0002$) (see Figure 2). The analysis of individual bins between 30 and 200 ms showed that the percentage of phonated intervals in the 30–100-ms bin was significantly less during Reading-1.0 s for the control group ($t = 2.53$) and the PWS ($t = 2.31$) at the .05 alpha level (see Table 4). Eight out of 10 controls, and 9 out of 10 PWS, produced a lower percentage of phonated intervals in the 30–100-ms bin during Reading-1.0 s than during control reading. The only other notable change was a reduction in the 101–150-ms bin for the PWS (effect size = $-.60$). All participants had a reduction in at least one of the three bins below 200 ms, with the majority (17) having a drop in 2 of the bins, 9 in the control group and 8 of the PWS. The *t*-tests examining the difference in the change from the control condition to Reading-1.0 s between the two groups for the 30–100-ms, 101–150-ms, and 151–200-ms bins were not significant. Finally, all PWS either increased mean vowel duration by 20% or more ($n = 5$), decreased the percentage of 30–100-ms phonated intervals by 30% or more ($n = 7$), or both ($n = 2$), during Reading-1.0 s

3.7. Association Between Speech Rate and Vocalization Variables

Pearson correlation coefficients were performed between the change in articulation rate from control reading to Reading-1.0 s and the change in specific vocalization variables (variables that changed significantly between conditions) from control reading to Reading-1.0 s. Two PWS had no 5-s stutterfree intervals during any of their trials, so these two participants were excluded from this analysis. A negative correlation indicates that greater reductions in articulation rate from control reading to Reading-1.0 s were associated with greater increases in the length or amount of the speech production variable. Change in articulation rate and change in vowel duration were significantly correlated for the controls ($r = -.92$, $p < .0001$) and for PWS ($r = -.98$, $p < .0001$). There was no significant correlation between change in articulation rate and change in the percentage of phonated intervals in the 30–100-ms bin for the controls ($r = -.05$) or for the PWS ($r = -.33$), nor was there a significant correlation between change in articulation rate and change in mean phonated interval length for the control group ($r = -.62$).

3.8. Association between speech naturalness ratings and speech production variables

Three significant correlations were found between speech production variables and speech naturalness during Reading-1.0 s. Vowel duration ($r = .75$, $p = .013$), mean phonated interval duration ($r = .64$, $p = .046$), and percentage of 101–150-ms phonated intervals ($r = -.79$, $p = .006$) were significantly correlated with speech naturalness ratings for PWS. There were no significant correlations for the controls.

4. Discussion

The primary research question in the present study was whether or not there were changes in speech production variables from a control condition to a rhythmic condition which involved alternating between 1 s of reading and 1 s of silence (Reading-1.0 s) for PWS and/or normally fluent controls. The results showed that mean vowel durations increased

significantly for both groups, mean phonated interval duration increased significantly for the control group, and the percentage of short (30–100 ms) phonated intervals decreased significantly for both groups. It is interesting that only controls showed a significantly longer mean phonated interval duration, as the other significant findings occurred for both groups. This was likely due to the larger (over 2 times greater) standard deviation for the PWS during control reading. Standard deviations for both groups during Reading-1.0 s were virtually identical. Finding greater variability in voiced segments during control conditions for a group of PWS compared to a control group is not unusual (e.g., Adams & Ramig, 1980; Janssen & Wieneke, 1987), and this was also found for vowel duration in the present study.

Other research questions in the present study included: (a) How will a group of naïve listeners rate the naturalness of the speech produced during Reading-1.0 s?; and (b) Which speech production variables are associated with naturalness ratings during Reading-1.0 s? Average speech naturalness ratings by the naïve listeners were approximately 7 on the 1–9 scale. Lastly, mean vowel duration and mean phonated interval duration were positively correlated, and the percentage of 101–150-ms phonated intervals was negatively correlated, with speech naturalness for the PWS.

The speech production changes suggest that extended vowel durations and a reduction in the percentage of short phonated intervals may be a part of the fluency-inducing mechanism of the Reading-1.0 s condition. The necessity of these changes for fluency during the experimental condition remains in question, however, for two reasons. First, changes in vowel duration were highly correlated with changes in articulation rate, and if speech rates were matched between conditions, vowel duration may not have been altered (Davidow et al., 2010). In fact, for the two participants (one control and one PWS) who had the closest match in articulation rates between control reading and Reading-1.0 s (7 SPM and 5 SPM difference, respectively), mean vowel duration actually decreased.

Second, both changes also occurred for the control group, who were fluent during both conditions, so it is possible that changes may simply be a byproduct of the condition rather than a change that is necessary for fluency. However, it is also possible that vocal fold vibration duration is altered similarly in both groups, and the adjustment still contributes to the fluency for PWS. Regardless of the necessity of the changes, they may be sufficient for PWS to receive fluency benefits from these adjustments. In addition, these results raise the interesting possibility that, even if a specific change in the manner of speech production is not found to be necessary, it could be beneficial to combine two fluency inducers (e.g., reduction in the percentage of short phonated intervals and the rhythmic pattern) to produce a powerful stuttering reduction procedure. The fact that speakers are naturally making speech production changes means that less learning may be necessary (than having to introduce the change) to take advantage of these fluency inducers.

4.1. Previous Metronome Effect Studies

Previous studies of the metronome effect using instatement styles of one syllable or one word per metronome beat have also found a reduction in the percentage of phonated intervals (Davidow et al., 2009) and increases in vowel durations (Brayton & Conture, 1978; Klich & May, 1982; Stager et al., 1997), suggesting that PWS might use these vocalization changes similarly to induce fluency during different rhythmic speaking styles. The reduction in percentage of short phonated intervals also occurred during singing, chorus reading, and prolonged speech in the Davidow et al. study. This reduction has now occurred in every FIC tested and has been shown to induce fluency when purposefully produced (Gow & Ingham, 1992; Ingham, Kilgo et al., 2001; Ingham et al., 1983). The present study's findings further support the possible importance of a reduction in short phonated intervals for fluency, and

future research should examine phonated interval distributions during less powerful FICs (auditory masking, slowed speech, the adaptation effect) to determine their possible contribution to those FICs.

Other vocalization variables that were altered in previous metronome effect studies and that were collected in the present study included intraoral peak pressure and intraoral pressure rise time. Decreased intraoral peak pressure (Hutchinson & Navarre, 1977; Stager et al., 1997) and increased intraoral pressure rise time (Stager et al., 1997) have been reported in previous studies, but neither of these variables changed significantly in the present study.

4.2. Speech naturalness

The naturalness ratings of approximately 7 on the 1–9 scale during Reading-1.0 s would be considered an inappropriate treatment outcome. This value is higher than that found by Finn and Ingham (1994), who reported mean ratings of approximately 5 for the same experimental condition. A likely reason for this difference is that the PWS rated their own speech in the Finn and Ingham study; it may be the case that rhythmic speech is perceived as more unnatural by listeners than speakers believe. Previous studies have found differences in naturalness ratings between listeners who stutter and listeners who do not stutter, with some reporting listeners who stutter as being less critical (see Teshima et al., 2010, for a review). This experiment also revealed a possible solution to this problem, however, because vowel duration and mean phonated interval duration were positively correlated, and the percentage of 101–150-ms phonated intervals was negatively correlated, with speech naturalness for the PWS during Reading-1.0 s. These are important findings, because if these variables are vital to the perception of speech naturalness during this condition, then they can and should be purposefully altered during the experimental condition to improve naturalness. Vowel duration will be interesting to study as well, because speakers may be lengthening it to promote fluency, but may need to shorten it for improved naturalness if it is found to be a critical factor in listener-rated naturalness. Other factors impacting speech naturalness may have been articulation rate (Logan, Roberts, Pretto, & Morey, 2002), position in the sentence that the speaker pauses, and stimulus-off duration length. Several judges mentioned these as factors in their ratings.

4.3. Future Research

Future research should examine the necessity of the vocalization changes found in the present study. As mentioned previously, the finding that the changes occurred in both groups limits the conclusions that can be drawn from the present study in that regard. One way to further examine if the vocalization changes are necessary for fluency is to have PWS produce the speech pattern while attempting to not alter whichever variable that changed when they were allowed to choose their own speech pattern. For example, if Participant 1 produced longer vowel durations during Reading-1.0 s in the present study, he could be instructed to try to shorten his vowel durations while performing the task again. This type of investigation is needed to determine the necessity of vocalization changes, something that previous vocalization and FIC studies (including the present study), which are mainly descriptive, have not examined. We are presently conducting these types of studies in our laboratory.

As mentioned previously, Jones and Azrin (1969) and Finn and Ingham (1994) found zero or near-zero levels of stuttering using stimulus-on durations of 0.1 s, 0.5 s, and 1 s, all coupled with a stimulus-off duration of 1 s. Since neither Jones and Azrin nor Finn and Ingham varied the stimulus-off duration, it would be interesting to examine stuttering frequency, acoustic and aerodynamic variables, and speech naturalness using stimulus-off durations other than 1.0 s. A combination of a 2.0-s stimulus-on duration and a 0.5-s

stimulus-off duration, for example, may have similar fluency benefits as the present study's experimental condition, yet with improved naturalness (naturalness improved with a longer stimulus-on duration in previous studies using RSLSD). The range of stimulus-off durations (no reading) for PWS in the present study was 0.77 – 1.17 s, which suggests that the stimulus-off duration does not need to be 1 s to eliminate stuttering. Examination of vocalization changes while varying the stimulus-off duration would also allow for further information regarding how this type of speaking style alters the speech mechanism.

How natural this variation of metronomic speech feels to the speaker, which was not examined in the present study (Finn & Ingham, 1994 reported feel naturalness ratings around 5), may also improve with a different stimulus-off duration, because pause duration varies as a function of several factors (e.g., speaker, discourse type, speech rate, and prosodic structure; see Krivokapic, 2007 for a review). Additionally, several participants in Finn and Ingham's study reported that they used "cognitive effort" and "attention to speech" as parameters for rating how natural their speech felt. A stimulus-off duration more closely aligned with a speaker's habitual mean pause duration may increase the feeling of naturally-produced speech or be perceived as less effortful. Lastly, speech naturalness during Reading-1.0 s might be improved if the speaker is provided with regular listener-rated naturalness feedback, as Ingham, Sato, et al. (2001) demonstrated during rhythmic speech with 0.1-s stimulus beats. Overall, these results suggest that a study combining a variety of stimulus-on durations, a variety of stimulus-off durations, and instructions for the participants to strive for natural speech (with accompanying biofeedback or listener-generated feedback) could provide valuable insight into the use of RSLSD as a viable treatment agent for stuttering.

Acknowledgments

This research was supported in part by grants from the National Institutes of Health, the American Speech-Language-Hearing Association (Advancing Academic-Research Career Award), and Hofstra University awarded to the first author. We would like to thank the participants for their time and the judges for all of their hard work during data analysis.

References

- Adams MR, Ramig P. Vocal characteristics of normal speakers and stutterers during choral reading. *Journal of Speech and Hearing Research*. 1980; 23:457–469. [PubMed: 7442204]
- Alm PA. Stuttering and the basal ganglia circuits: a critical review of possible relations. *Journal of Communication Disorders*. 2004; 37:325–369. [PubMed: 15159193]
- Andrews G, Guitar B, Howie PM. Meta-analysis of the effects of stuttering treatment. *Journal of Speech and Hearing Disorders*. 1980; 45:287–307. [PubMed: 7412224]
- Andrews G, Howie PM, Dozsa M, Guitar BE. Stuttering: Speech pattern characteristics under fluency-inducing conditions. *Journal of Speech and Hearing Research*. 1982; 25:208–216. [PubMed: 7120960]
- Barber V. Studies in the psychology of stuttering, XVI: Rhythm as a distraction in stuttering. *Journal of Speech Disorders*. 1940; 5:29–42.
- Barlow, SM.; Suing, G.; Andreatta, RD. Speech aerodynamics using AEROWIN. In: Barlow, SM., editor. *Handbook of clinical speech physiology*. San Diego: Singular; 1999. p. 165-189.
- Bothe AK, Davidow JH, Bramlett RE, Ingham RJ. Stuttering treatment research 1970–2005:I. Systematic review incorporating trial quality assessment of behavioral, cognitive, and related approaches. *American Journal of Speech-Language Pathology*. 2006; 15:321–341. [PubMed: 17102144]
- Brady JP. Studies on the metronome effect on stuttering. *Behavioral Research and Therapy*. 1969; 7:197–204.

- Brady JP. Metronome-conditioned speech retraining for stuttering. *Behavior Therapy*. 1971; 2:129–150.
- Brayton ER, Conture EG. Effects of noise and rhythmic stimulation on the speech of stutterers. *Journal of Speech and Hearing Research*. 1978; 21:285–294.
- Davidow JH, Bothe AK, Andreatta RD, Ye J. Measurement of phonated interval during four fluency-inducing conditions. *Journal of Speech, Language, and Hearing Research*. 2009; 52:188–205.
- Davidow JH, Bothe AK, Richardson JD, Andreatta RD. Systematic studies of modified vocalization: Effects of speech rate and instatement style during metronome stimulation. *Journal of Speech, Language, and Hearing Research*. 2010; 53:1579–1594.
- Davis, CS. *Statistical Methods for the Analysis of Repeated Measurements*. New York: Springer-Verlag; 2002.
- Finn P, Ingham RJ. Stutterers' self-ratings of how natural speech sounds and feels. *Journal of Speech and Hearing Research*. 1994; 37:326–340. [PubMed: 8028313]
- Gow ML, Ingham RJ. Modifying electroglottograph-identified intervals of phonation: The effect on stuttering. *Journal of Speech and Hearing Research*. 1992; 35:495–511. [PubMed: 1608241]
- Greenberg JB. The effect of a metronome on the speech of young stutterers. *Behavior Therapy*. 1970; 1:240–244.
- Hammen VL, Yorkston KM. Speech and pause characteristics following speech rate reduction in hypokinetic dysarthria. *Journal of Communication Disorders*. 1996; 29:429–445. [PubMed: 8956101]
- Hochberg, Y.; Tamhane, AC. *Multiple comparison procedures*. New York: Wiley; 1987.
- Hutchinson JM, Navarre BM. The effect of metronome pacing on selected aerodynamic patterns of stuttered speech: Some preliminary observations and interpretations. *Journal of Fluency Disorders*. 1977; 2:189–204.
- Ingham, RJ.; Bakker, K.; Ingham, JC.; Moglia, R.; Kilgo, M. *Stuttering Measurement System (SMS)*. Santa Barbara, CA: University of California, Santa Barbara; 2005. (available at <http://speech.ucsb.edu/>)
- Ingham, JC.; Ingham, RJ. *The stuttering measurement system (SMS) training manual*. Santa Barbara, CA: University of California, Santa Barbara; 2004.
- Ingham RJ, Kilgo M, Ingham JC, Moglia R, Belknap H, Sanchez T. Evaluation of a stuttering treatment based on reduction of short phonation intervals. *Journal of Speech, Language, and Hearing Research*. 2001; 44:1229–1244.
- Ingham, R.; Moglia, R.; Kilgo, M.; Felino, A. *Modifying phonation intervals stuttering treatment program: Training and evaluation system, version 1.8.0*. University of California; Santa Barbara: 2006.
- Ingham RJ, Montgomery J, Ulliana L. The effect of manipulating phonation duration on stuttering. *Journal of Speech and Hearing Research*. 1983; 26:579–587. [PubMed: 6668945]
- Ingham RJ, Sato W, Finn P, Belknap H. The modification of speech naturalness during rhythmic stimulation treatment of stuttering. *Journal of Speech, Language, and Hearing Research*. 2001; 44:841–852.
- Janssen, P.; Wieneke, G. The effects of fluency inducing conditions on the variability in the duration of laryngeal movements during stutterers' fluent speech. In: Peters, HFM.; Hulstijn, W., editors. *Speech motor dynamics in stuttering*. New York: Springer-Verlag/Wien; 1987. p. 337–344.
- Johnson W, Rosen L. Studies in the psychology of stuttering: VII: Effect of certain changes in speech pattern upon frequency of stuttering. *Journal of Speech Disorders*. 1937; 2:105–110.
- Jones RJ, Azrin NH. Behavioral engineering: Stuttering as a function of stimulus duration during speech synchronization. *Journal of Applied Behavior Analysis*. 1969; 2:223–229. [PubMed: 16795224]
- Klich RJ, May GM. Spectrographic study of vowels in stutterers' fluent speech. *Journal of Speech and Hearing Research*. 1982; 25:364–370. [PubMed: 7176608]
- Krivokapic J. Prosodic planning: Effects of phrasal length and complexity on pause duration. *Journal of Phonetics*. 2007; 35:162–179. [PubMed: 18379639]

- Logan KJ, Roberts RR, Pretto AP, Morey MJ. Speaking slowly: Effects of four self-guided training approaches on adults' speech rate and naturalness. *American Journal of Speech-Language Pathology*. 2002; 11:163–174.
- Maassen B, Povel DJ. The effect of correcting temporal structure on the intelligibility of deaf speech. *Speech Communication*. 1984; 3:123–135.
- Mallard AR, Westbrook JB. Vowel duration in stutterers participating in precision fluency shaping. *Journal of Fluency Disorders*. 1985; 10:221–228.
- Martin RR, Haroldson SK. Stuttering identification: Standard definition and moment of stuttering. *Journal of Speech and Hearing Research*. 1981; 24:59–63. [PubMed: 7253630]
- Martin RR, Haroldson SK. Stuttering and speech naturalness: Audio and audiovisual judgments. *Journal of Speech and Hearing Research*. 1992; 35:521–528. [PubMed: 1376781]
- Martin RR, Haroldson SK, Triden K. Stuttering and speech naturalness. *Journal of Speech and Hearing Disorders*. 1984; 49:53–58. [PubMed: 6700202]
- Metz DE, Schiavetti N, Sacco PR. Acoustic and psychophysical dimensions of the perceived speech naturalness of nonstutterers and posttreatment stutterers. *Journal of Speech and Hearing Disorders*. 1990; 55:516–525. [PubMed: 2381193]
- Nakagawa S. A farewell to Bonferroni: the problems of low statistical power and publication bias. *Behavioral Ecology*. 2004; 15:1044–1045.
- O'Brian S, Onslow M, Cream A, Packman A. The Camperdown program: Outcomes of a new prolonged-speech treatment model. *Journal of Speech, Language, and Hearing Research*. 2003; 46:933–946.
- Onslow M, Adams R, Ingham R. Reliability of speech naturalness ratings of stuttered speech during treatment. *Journal of Speech and Hearing Research*. 1992; 35:994–1001. [PubMed: 1447933]
- Öst L, Götestam G, Melin L. A controlled study of two behavioral methods in the treatment of stuttering. *Behavior Therapy*. 1976; 7:587–592.
- Schaeffer N, Eichor N E. The effects of differential vowel prolongations on perceptions of speech naturalness. *Journal of Fluency Disorders*. 2001; 26:335–348.
- Schiavetti, N.; Metz, DE. Stuttering and the measurement of speech naturalness. In: Curlee, RF.; Siegel, GM., editors. *Nature and treatment of stuttering: New directions*. 2. Boston: Allyn and Bacon; 1997. p. 398–412.
- Stager SV, Denman DW, Ludlow CL. Modifications in aerodynamic variables by persons who stutter under fluency-evoking conditions. *Journal of Speech, Language, and Hearing Research*. 1997; 40:832–847.
- Stager SV, Jeffries KJ, Braun AR. Common features of fluency-evoking conditions studied in stuttering subjects and controls: an H₂¹⁵O PET study. *Journal of Fluency Disorders*. 2003; 28:319–336. [PubMed: 14643068]
- Stager SV, Ludlow CL. Speech production changes under fluency-evoking conditions in nonstuttering speakers. *Journal of Speech, Language, and Hearing Research*. 1993; 36:245–253.
- Teshima S, Langevin M, Hagler P, Kully D. Post-treatment speech naturalness of Comprehensive Stuttering Program clients and differences in ratings among listener groups. *Journal of Fluency Disorders*. 2010; 35:44–58. [PubMed: 20412982]
- Toothaker, LE. *Multiple Comparisons for Researchers*. Newbury Park, CA: Sage Publications, Inc; 1991.
- Trajkovski N, Andrews C, Onslow M, Packman A, O'Brian, Menzies R. Using syllable-timed speech to treat preschool children who stutter: A multiple baseline experiment. *Journal of Fluency Disorders*. 2009; 34:1–10. [PubMed: 19500711]
- Van Riper, C. *The treatment of stuttering*. Englewood Cliffs, NJ: Prentice-Hall; 1973.
- Wechsler, D. *Wechsler Adult Intelligence Scale*. 3. San Antonio, TX: Harcourt Assessment; 1997. (WAIS-III)
- Wickens, TD. *Multiway contingency tables analysis for the social sciences*. Hillsdale, NJ: Lawrence Erlbaum; 1989.
- Wingate ME. Sound and pattern in “artificial” fluency. *Journal of Speech and Hearing Research*. 1969; 12:677–686. [PubMed: 5385718]

Wingate ME. Effect on stuttering of changes in audition. *Journal of Speech and Hearing Research*. 1970; 13:861–873. [PubMed: 5491360]

Biographies

Jason H. Davidow, Ph.D., is an Assistant Professor in the Speech-Language-Hearing Sciences Department at Hofstra University. His main research interests include stuttering treatment outcome and the measurement of speech production changes during fluency-inducing conditions in persons who stutter.

Anne K. Bothe, Ph.D., CCC-SLP, is a Professor in The Communication Sciences and Special Education Department at the University of Georgia. Her research and writing focus on the intersection of measurement and treatment variables for stuttering.

Jun Ye, Ph.D., is an Assistant Professor of Biostatistics at South Dakota State University (SDSU) and is also a statistical consultant at SDSU Agriculture Experiment Station. He also worked as a Biostatistician and Epidemiologist for the Department of Medicine at the Massachusetts General Hospital until August 2010.

Appendix

Carrier Phrases Used for Aerodynamic Data

1. The *pitifully* sad guard saw *Bobby* in the *basement*.
2. Wendy *panicked* when the *boozier* fell over *Pablo*.
3. My *bulging* arms grow *bigger* and very *powerful*.
4. She saw *Peter* on the new *better* and high *podium*.
5. Her *biting* dog Snow *painfully* began *chewing*.
6. Toy *batteries* cost a few *pennies* from my *bookkeeper's* store.

Carrier Phrases Used for Acoustic Data

1. Those ladies, near the fence, are *taking* the *paper sacks* to the barn.
2. He seems *thoughtful* to spend a lot of time and help *search* the house.
3. The *people* in the corner of the yard look like *puppets faithfully* working.
4. The nicely dressed man is *touching* the *shirt* while *checking* out the very beautiful view.
5. If you look closely, you can see a *calf* on the gigantic hill in the back of the *picture*.
6. That small boy is *keeping* a watchful eye on the large, wooden treasure bucket.
7. The girls with the extremely long hair are *shockingly* surprised at how he *fight*s.
8. The *sassy* ladies appear to be waiting *patiently* and want to *attack* the valuable treasure.
9. Nobody is *speaking*, as all of the *spectators* really just want to *focus* on the show.
10. They may be saying that he *put* on a *spectacular* show for the *special* occasion.

Research Highlights

- Vowel duration increased during variation of metronomic speech
- Percent of short phonated intervals reduced during variation of metronomic speech
- Vowel duration correlated with speech naturalness in Persons who stutter
- Mean phonated interval duration correlated with speech naturalness in Persons who stutter
- Percent of short phonated intervals correlated with speech naturalness in Persons who stutter

She saw peter on the new better and high podium.

Figure 1.

An example of the aerodynamic signals from the AEROWIN (Neuro Logic, Inc., Lawrence, KS) display for one carrier phrase during control reading. The top channel is the audio trace, which is displayed in volts. The middle channel is air flow, which is displayed in cubic centimeters per second (cc/sec). The bottom channel is intraoral pressure, which is displayed in centimeters of water (cmH₂O).

Figure 2. Distribution of phonated intervals for the control reading versus Reading-1.0 s comparison. Data are the average values across the group. Errors bars represent the standard deviation of the means.

Table 1

Example of the order of experimental tasks for the first day of the larger project. Each grouping of data-collecting trials (acoustic and aerodynamic) was preceded by 30-s practice trials.

Day 1
Four 3-min monologues
Four 3-min readings
Control reading: 30-s aerodynamic trials, 75-s acoustic trials
Condition X: 75-s acoustic trials, 30-s aerodynamic trials
Rest period (2-min monologues)
Reading-1.0 s: 30-s aerodynamic trials, 75-s acoustic trials
Rest period (2-min monologues)
Condition X: 75-s acoustic trials, 30-s aerodynamic trials
Rest period (2-min monologues)
Condition X: 30-s aerodynamic trials, 75-s acoustic trials
Rest period (2-min monologues)
Condition X: 75-s acoustic trials, 30-s aerodynamic trials

Note. The order of the aerodynamic and acoustic trials was reversed in each subsequent condition to conserve time by not needing to remove and replace equipment (microphone and accelerometer) as frequently. Condition X refers to the conditions not analyzed for the present study.

Table 2

The mean (using absolute values) and range of differences between measurement occasions, and percent deviation scores for interjudge and intrajudge reliability data for all dependent variables.

Reliability	Variable									
	Peak pressure(cm H ₂ O)	Pressure rise time(ms)	Maximum air flow(cc/s)	Vowel midpoint airflow (cc/s)	Vowel duration(ms)	Voice onset time (ms)	Fundamental frequency(Hz)	Syllables per minute	Stimulus-on & stimulus-off durations: on off	
Interjudge										
Mean difference	0.00	5.50	.05	16.19	9.13	5.48	1.40	13.11	.02	.02
Range of differences	0.00-0.01	0.00-61.30	0.00-2.08	0.00-65.94	0.18-62.74	0.15-31.34	0.01-26.56	0.00-37.33	0-.19	0-.18
Percent Deviation	0.00	5.58	0.02	12.79	7.74	10.52	1.07	6.49	1.42	2.23
Intrajudge										
Mean difference	0.00	1.67	0.00	6.43	2.35	1.47	.88	8.51	.01	.01
Range of differences	0.00-0.00	0.00-29.14	0.00-0.08	0.00-74.15	0.00-19.13	0.00-10.58	0.00-16.35	0.02-32.10	0-.09	0-.08
Percent Deviation	0.00	1.24	0.00	4.21	2.52	3.16	0.60	3.93	1.30	1.34

Table 3

Mean percent syllables stuttered (%SS), syllables per minute (SPM), articulation rate (displayed in SPM), and naturalness ratings for speaking tasks (standard deviations in parentheses).

Dependent Variable	Speaking Condition			
	12 minutes of reading	12 minutes of monologue	Control Reading	Read-1.0 s
Persons Who Stutter				
%SS	5.55 (9.15)	7.13 (7.83)	5.19 (7.59)	0.28 (0.78)
SPM	217.73 (84.73)	175.15 (56.44)	226.47 (65.05)	138.33 (23.21)
Stutter-free articulation rate	--	--	306.42 (41.51)	246.67 (53.74)
Naturalness	--	--	3.48 (2.50)	6.89 (.48)
Control Group				
SPM	--	238.95 (21.87)	298.99 (13.82)	148.04 (7.65)
Articulation rate	--	--	312.66 (26.72)	261.52 (40.09)
Naturalness	--	--	1.54 (.37)	7.03 (.76)

Note. -- = data not analyzed for task.

Table 4

Mean values (with standard deviation in parentheses) and effect sizes (Cohen's d) for each dependent variable during both speaking conditions.

Dependent Variable	Control group			Persons who stutter		
	Effect size	Experimental mean	Control mean	Effect size	Experimental mean	Control mean
Peak pressure (cm H ₂ O)	-.59	6.95 (1.23)	7.88 (1.86)	.22	7.13 (2.09)	6.66 (2.22)
Pressure rise time (ms)	-.31	83.50 (20.35)	89.91 (21.50)	.31	96.23 (28.16)	89.03 (15.94)
Maximum air flow(cc/s)	-.43	551.22 (193.65)	662.71 (312.13)	-.17	551.73 (232.73)	597.95 (314.33)
Vowel midpoint air flow (cc/s)	-.24	149.29 (60.57)	165.63 (75.25)	-.06	148.91 (70.14)	153.51 (75.12)
Vowel duration (ms)	.99*	132.15 (24.81)	113.67 (8.86)	.72*	154.61 (34.09)	130.99 (31.39)
Voice onset time (ms)	.10	55.66 (10.67)	54.72 (8.85)	.30	58.22 (15.61)	53.61 (15.56)
Fundamental frequency (Hz)	.07	149.78 (38.78)	147.20 (30.97)	.11	132.66 (35.09)	128.59 (35.41)
Phonated intervals						
Overall mean (ms)	.80**	303.98 (74.15)	256.51 (39.31)	.40	330.43 (74.02)	296.63 (92.07)
30–100–ms bin	-.69*	12.04 (8.38)	16.80 (5.13)	-.80*	9.35 (5.09)	17.38 (13.27)
101–150–ms bin	-.41	14.49 (6.92)	16.63 (2.49)	-.60	12.66 (6.14)	15.81 (4.24)
151–200–ms bin	-.21	13.07 (2.14)	13.59 (2.86)	.01	15.20 (4.78)	15.17 (3.06)

* $p < .05$.

** $p = .01$.