4-26-2016

Acoustic Correlates of Speech Naturalness in Post-Treatment Adults Who Stutter: Role of Fundamental Frequency

Jacqueline Schmitt
Western Michigan University, jacquelineschmitt@gmail.com

Acoustic Correlates of Speech Naturalness in Post-Treatment Adults Who Stutter: Role of Fundamental Frequency

Jacqueline Schmitt

Western Michigan University

Abstract

Purpose: Stuttering therapy programs can be effective in reducing the frequency of stuttering events. However, one complication of some stuttering treatment programs is that speech patterns following treatment can sound unnatural to listeners. While altered intonation patterns have been implicated as possible cause of unnatural sounding speech, such a relationship has not been empirically examined. The goal of the current study is to investigate whether measures of speaking fundamental frequency (SFF) are associated with listener ratings of speech naturalness in adults who stutter following participation in a stuttering treatment program.

Methods: Twenty nine adults who stutter and just completed an intensive stuttering treatment program were drawn from a larger data base of adults who stutter and normally fluent speakers. Measures of mean speaking fundamental frequency (SFF) and SFF variability were made from audio recordings of oral reading samples. These measures were correlated with previously collected listener ratings of speech naturalness.

Results: Mean SFF and SFF variation were not significantly correlated with post-treatment naturalness ratings. A negative correlation was observed between post-treatment stuttering severity scores and measures of SFF variability indicating that those with the lowest stuttering severity measures were associated with largest SFF variability.

Conclusions: Measures of speaking fundamental frequency did not correlate with listener's ratings of speech naturalness. It is possible the SFF variability measures used in the current study were too coarse to reflect any relevant anomalies in the intonation patterns present in the speech samples.

Background

Stuttering, a speech disorder that typically begins in early childhood, is characterized by involuntary disruptions in speech fluency. It often presents as part or whole word repetitions (bu-bu-but or and-and-and), prolongations (Mo—mmy), or blocks of airflow before or during a sound or word production (Yairi & Seery, 2015). Prevalence of this complex speech disorder is approximately 1% of the general population (Yairi & Seery, 2015), affecting approximately 70 million people worldwide.

A common approach to treating stuttering involves altering speech patterns in order to reduce the likelihood of stuttering. These treatments can focus on a variety of targets including increased abdominal breathing, adequate breath support and volume, creating a continuity of airflow before the initiation of voicing, producing smooth articulatory and phonatory onsets, maintaining continuous phonation throughout the speech breath, and learning to properly time speech breaths within the utterance. These targets are practiced and speech is "shaped" until they can be generalized into a client's everyday speech (Yairi & Seery, 2015). A possible side effect of altering the entire manner of speaking is that, following intervention, participants' speech may be perceived by themselves and others as sounding less natural (Martin, Triden, & Haroldon, 1984).

Fluent, but less natural speech is considered a possible negative therapy outcome for both the clinician delivering treatment and the client receiving services. The intent of therapy is to aid in reducing stuttering, but not to exchange one form of perceived change in speech naturalness for another. To measure these changes, one must ask how to quantify perceptions of speech naturalness. Perhaps one reason the role of naturalness or normalcy in the treatment process for

persons who stutter has not been studied extensively or systematically is the lack of a convenient means for naturalness perception quantification (Martin et al, 1984).

Measuring Speech Naturalness

An initial investigation of measuring speech naturalness in persons who stutter was by Martin and colleagues (1984). This study set out to determine if speech naturalness was a reliable, scalable dimension of speech in persons who stutter that could be used in future studies. Martin and colleagues gathered speech samples from 10 adults who stutter, 10 adults who do not stutter, and 10 adults who stutter while speaking with a 250ms delay in auditory feedback (DAF). DAF was utilized as it is considered to enhance speech fluency. One minute segments without perceptible stuttering were selected from each sample and played for 30 undergraduate students. The listeners rated the naturalness of each segment on a scale from 1 (highly natural) to 9 (least natural) based on their auditory perceptual impressions.

The data revealed that the mean naturalness rating assigned to the samples of the adults who stuttered were significantly higher or rated less natural than the mean rating assigned to the samples of the adults who did not stutter. The mean rating assigned the DAF samples from the adults who stutter was also significantly higher, reflecting less natural sounding speech, than the mean rating for the samples from the adults who did not stutter (Martin et al, 1984). To assess the consistency of this method of naturalness scaling, each sample rating was compared to each of the other ratings completed by all 30 student raters. These comparisons revealed raters did not differ by more than plus or minus one scale value for 75% of the assigned values. The high interrater agreement suggested that observers can consistently attend to and quantify speech naturalness by utilizing a 1-9 naturalness scale. Martin and colleagues concluded that, if future

research correlated with their findings, speech naturalness could be used in the study of persons who stutter as a measure of outcome.

Metz, Schiavetti, and Sacco (1990) focused on treated adults who stutter to further test the validity of the 1-9 naturalness scaling procedures proposed by Martin and colleagues. Thirty undergraduate students evaluated audio segments of speech from 20 persons who stutter following participation in a gentle voicing treatment program. The speech segments were comprised of both descriptions of a picture and recitations of a reading passage. The students were not given a definition of naturalness but were asked to rate both sample types on the 1-9 naturalness interval scale and using a direct magnitude rating procedure. A strong linear correlation was found between the two methods suggesting the two approaches were comparable for both reading paragraphs and the picture descriptions.

Previous studies' have demonstrated that speech naturalness is a scalable dimension in the speech of persons who stutter both before and after treatment. These studies identify a decrease in speech naturalness after treatment. This leaves the question: what causes less natural ratings of speech of adults who stutter following stuttering treatment?

Acoustic Correlates of Speech Naturalness

In addition to assessing the psychometric properties of naturalness rating scales Metz et al. (1990), also assessed the relationship between selected speech acoustic measures and perceptual measures of speech naturalness. Regression was used to determine associations between speech naturalness ratings (for both the picture description and paragraph reading) and selected acoustic measures. Two factors significantly predicted speech naturalness ratings. For the picture description task, speech naturalness was highly correlated with voice onset time (VOT) (r=.679), which is a measure of the time between the burst release of a stop consonant

and the onset of voicing. For the paragraph reading, speech naturalness was significantly predicted by sentence duration (r=.637) (Metz et al, 1990). Though not statistically significant, pauses between sentences and vowel duration measures were identified as other potentially influential predictors of speech naturalness. Although they identified these possible predictors of naturalness, the authors called for future studies to identify the extent of these characteristics' effects on naturalness in hopes to target such factors in clinical work.

Armson and Kiefte (2008) examined speech naturalness in a group of 31 adults who stuttered. All but one were receiving behavioral therapy. Participants were recorded within two contexts, producing a monologue a reading aloud passage under conditions, altered and unaltered auditory feedback. Thirty adult volunteers were recruited to evaluate speech naturalness of the samples according to the 1-9 interval rating scale utilized by Martin et al. (1984). Interval scale ratings were provided for 14 of the 31 participants, in both reading and monologue tasks, by 15 raters.

For both reading and monologue tasks, mean naturalness ratings were markedly better during the altered feedback condition. Additionally, several participants received poor naturalness scores in that context despite low amounts of stuttering which led Armson and Kiefte (2008) to believe both participants' and listeners' perceptions of speech quality had been influenced by additional acoustical variables, other than speech rate, and overt instances of stuttering.

Schaeffer and Eichorn (2001) examined the relationship between vowel prolongation and listener perception of speech naturalness. A resynthesis technique was used to systematically prolong the vowel in twelve monosyllable words contained within a carrier phrase. All were produced by the same speaker. Vowels were prolonged so they were 150%, 200%, and 250%

longer than the original vowel duration. Samples were presented to 52 college students who rated each sample's naturalness. Listeners recorded whether they "strongly agreed," "agreed," or "disagreed" that a speech sounded natural. During data analysis, the ratings for "strongly agree" and "agree" were combined, because both ratings were perceived as natural (Schaeffer & Eichorn, 2001). This resulted in a dichotomous "natural" vs. "unnatural" rating system. Results indicated that 83% of the participants perceived samples to be natural at 150% of the original duration, about 50% of the participants perceived samples to be natural at 200% of the original duration and and very few listeners perceived the samples that were 250% of original duration to be natural (Schaeffer & Eichorn, 2001). These results indicated that vowel duration does affect perceived speech naturalness and that naturalness ratings are strongly influenced by the extent of vowel prolongation. While this study did not specifically examine speech of person who stutters, the results suggest that sound prolongation, which is part of many stuttering treatment programs, has a marked influence on speech naturalness perception.

Schaeffer (2006) followed up with a study that examined the relation between sentence duration and listener ratings of naturalness. Twelve sentence samples were recorded from a fluent speaker and each phrase was prolonged to 150%, 170%, 190%, and 210% the original duration. Listeners were 60 college students. Speech naturalness was rated using the methods of Schaffer and Eichorn (2001).

Analyses revealed that 100% of the participants perceived the original phrases as natural; 85% of the participants rated the phrases as natural at the 150% of the original duration; 54% of the participants perceived the phrases to be natural at 170% of the original duration; 24% perceived the phrases to be natural at the 190% of the original duration; and only 10% of the participants perceived the phrases to be natural at the 210% of the original duration (Schaeffer,

2006). Overall, the study found that, like vowel duration, varying phrase duration influenced listeners' perceptions of speech naturalness. A comparison of the data in this Schaeffer (2006) follow up study and the original study by Schaffer and Eichorn (2001) revealed that, although prolongation negatively affects speech naturalness, speech naturalness can be maintained if the (level of vowel and phrase) prolongation is not too extreme.

Overall, these studies provide some evidence for acoustic correlates of speech naturalness in stuttering and non-stuttered speech. To date, acoustic correlates of speech naturalness consist mainly of durational measures (sentence and phrase duration, vowel duration, and VOT duration). However, these measures do not fully account for the variance in speech naturalness ratings suggesting that other acoustic variables may play a role. For example, it is not uncommon to observe atypical intonation patterns in persons who stutter following stuttering treatment. However, there are no studies that directly examine the relationship between measures of speaking fundamental frequency (SFF) and perceived naturalness in individuals who stutter following treatment.

There have been studies examining SFF differences in persons who do and do not stutter. For example, Healey (1982) measured SFF in a group of 10 adults who stuttered and 10 who did not stutter. Speech samples were collected from the participants in the form of audio recordings of two sentence length utterances: one declarative and one interrogative constructed and embedded in the middle of two separate short phrases. Mean SFF, SFF standard deviation, and SFF range were derived for all fluent samples. SFF measures were found to be significantly different between the two groups for both sentence types. Overall, the SFF standard deviations for the adults who did not stutter were greater than for the adults who did not stutter during the production of the declarative and interrogative sentences (Healey, 1982). This suggests that the

adults who stuttered tended to produce a more restricted SFF range during fluent speech production. The researcher called for analysis of SFF in adults who stutter to be compared with that of other parameters that appear in the fluent speech of adults who stutter (Healey, 1982). Although the study did not compare naturalness ratings to acoustic characteristics of speech, it revealed changes in SFF as potentially important variable in the speech of adults who stutter.

In summary, previous studies have focused mostly on the relationship between durational measures of speech events and listener perceptions of speech naturalness. Data from a variety of sources support the general conclusion that increased duration in speech events at the phoneme, word and phrase level are associated with reduced speech naturalness. The impact of speaking fundamental frequency variability on speech naturalness has not been examined to the same extent as durational measures even though clinicians often note that unnatural sounding speech often has atypical intonation patterns.

Aim of Study

The aim of the current study is to determine if speaking fundamental frequency variation is associated with listener ratings of speech naturalness in group of adults who stuttering following participation in an intensive stuttering treatment program.

Method

Participants

Participants were drawn from the Walter Reed Army Medical Center- Western Michigan University Stuttering Database, a large physiological and behavioral dataset that includes speech samples from a total of 43 adults who stutter and 43 healthy adult controls. A total of 29 adults (27 M, 2 F) with a history of stuttering since childhood served as participants. The participant group had a mean age of 25 years and 8 months with a range of 19-43 years. Originally, there were 34 participants whose datasets contained both pre- and post-treatment recordings of reading and monologue tasks. Five of these samples were removed from this study upon their initial review; three were omitted due to lack of post-treatment naturalness ratings, one was removed due to poor audio quality of reading sample, and one was removed because the participant's foreign accent was believed to be strong enough to have affected naturalness ratings.

Treatment Program

All participants were enrolled in and completed the Walter Reed Stuttering Treatment Program. The Walter Reed Stuttering Treatment Program was a 1-month, group-based fluency-shaping program comprised of three parts: Part one was instructional and addressed the nature of speech production, individual characteristics of stuttering, and attitudes and feelings associated with stuttering. Part two involved establishing behavioral targets including increased abdominal breathing, adequate breath support and volume, continuity of airflow, pre-voiced exhalation, easy articulatory and phonatory onset, continuous phonation, and proper use of phrasing. Part three of the program provided opportunities for participants to transfer fluency-enhancing skills from therapy to a variety of challenging speaking situations.

Sample Selection

Although the original naturalness ratings were derived from graduate student listener panel interpretations of one minute monologue clips of participants, the SFF examination in this study was based on analysis of the reading samples (see Appendix 1) that were collected in addition to the monologue speech samples. The reading samples were chosen for analysis due to the consistency in phonetic content and syllable count, making it easier to compare performance across each participant. The monologue clips were quite varied in regard to length, speech composure, content, and animation, and the consistency of the reading samples was more reasonable to utilize when considering future study reproducibility.

Severity Rating Procedure

The database includes information about each participant's stuttering severity before and after treatment. The SSI-3 was selected as a measure of stuttering severity for this study because it was judged by two speech-language pathologists to provide a valid continuous measure of the overall strength of the overt characteristics of stuttering. These two speech-language pathologists, who had experience with stuttered speech, served as judges and completed the SSI-3 measurement procedures. The measurements were conducted for the entirety of participant's original monologue video. Measurement procedures used a consensus format, meaning, when both judges agreed on an instance of stuttering or observed a physical concomitant, it was appropriately marked or was listened to and analyzed until an agreement on severity was reached. If an agreement was not reached, another judgment task was performed at least 6 months after the original scoring procedure. Once more, the videotape was watched and the judges discussed the sample until an agreement was reached.

Speech Naturalness Ratings

The database also includes speech naturalness ratings for many of the participants. Speech naturalness ratings were based on one minute samples that were randomly selected from the pre-treatment and post-treatment video recordings of each participant's monologue. Three different listener panels were then formed, each comprised from a pool of 20 graduate students enrolled in a speech and hearing program. Each panel rated a different randomized group of participants after having received identical instructions and samples presentations. Listeners were told to rate the naturalness of each sample along 9-point interval scales with 1 labeled as highly natural speech and 9 labeled as least natural speech.

Once the rating procedures were complete, the mean rating across the listener panel for each sample was calculated. The researchers then assessed for inter-rater reliability of the three groups by using the intraclass correlation coefficient (ICC), resulting in ICCs of 0.98, 0.97, and 0.95 for each of the three listener groups. These data were stored in the Walter Reed Army Medical Center- Western Michigan University Stuttering Database and the utilization of the participant information and listener ratings was approved by the Human Subjects Institutional Review Board at Western Michigan University and Walter Reed Army Medical Center.

Speaking Fundamental Frequency (SFF) Analysis

Speaking fundamental frequency (SFF) contours for the reading passage was generated using TF32, an acoustic analysis software package, TF32. The SFF contours were visually displayed along with the speech waveform. The interactive program allowed for editing of the samples. Editing occurred in cases where the fundamental frequency extraction algorithm identified values in non-voiced segments or when there were abrupt increases of decreases in pitch that did not align with the speech audio sample. Next, the SFF data were processed and

analyzed using a custom written Matlab routine. Any fundamental frequency values greater than three standard deviations from the mean were considered outliers and dropped from the data set. See Appendix 2 for step-by-step details regarding the speaking fundamental frequency extraction method.

The custom written Matlab program calculated a number of SFF measures based on the several thousand fundamental frequency measurements, derived for each speech sample. This included the mean SFF (Hz) and two measures of SFF variability: SFF standard deviation and SFF interquartile range. SFF variability measures were expressed both in Hertz (Hz) and in semitones (ST). Semitone calculations were made using the method described in Baken (1987).

Inter-rater reliability was examined by having a second undergraduate student majoring in Speech Pathology and Audiology perform the SFF editing task on three randomly selected participants. The results are in Table 1. Overall, reliability for the procedure appeared good with little variation in the results for three participants.

Table 1. Reliability comparisons of SFF mean, SFF standard deviation (Std) in hertz (Hz) and in semitones (ST), and interquartile range (IQR) in hertz and in semitones for speech samples from three participants following the pitch contour editing procedure.

Sample	Rater Initials	Mean (Hz)	Std (Hz)	Std (ST)	IQR (Hz)	IQR (ST)
Sample 1	CJ	106	13.73	2.09	15.96	2.60
Sample 1	JS	107	13.87	2.11	15.98	2.60
Sample 2	CJ	141	13.52	1.57	15.71	1.92
Sample 2	JS	141	13.63	1.58	15.87	1.94
Sample 3	CJ	111	15.98	2.31	17.86	2.76
Sample 3	JS	112	15.05	2.17	17.63	2.72

Results

Individual behavioral and acoustic data for all participants are summarized in Table 2. The pre-treatment stuttering severity scores ranged from 17-35 with a mean of 25.44. Post-treatment stuttering severity within the group ranged from 0-20 with a mean score of 9.20. Pretreatment naturalness ratings ranged from 1.20-7.20 with a mean rating of 4.39. Post-treatment naturalness ratings ranged from 1.65-6.41 with a mean rating of 3.20. Mean SFF ratings between the samples ranged from 102 – 208 Hz with a mean of 134 Hz. SFF standard deviation range of the samples were 11.16-31.97 Hz with a mean standard deviation of 17.80 Hz. When analyzing standard deviation of semitones, the range was 1.53-3.58 semitones with a mean of 2.16 semitones. The interquartile range across samples was 9.75-47.64 Hz with a mean of 22.41 Hz. The interquartile range in semitones varied from1.36 – 4.86 with a mean of 2.90 Hz.

Relationship Between Stuttering Severity and Speech Naturalness Ratings

Figure 1 is a plot showing the relationship between post-treatment stuttering severity and post-treatment speech naturalness. Speech naturalness is persons who stutter can be a function of the stuttering itself (which can be deemed unnatural) or due to other factors such as treatment related changes in speaking patterns. The relationship between post treatment stuttering severity measures and post treatment speech naturalness ratings was not significant (Figure 1) suggesting no systematic relationship between stuttering behaviors (as measured by the SSI-3) and speech naturalness ratings. Out of the 29 participants selected for this study, 30% of their naturalness scores worsened post-treatment, despite all having experienced reduction in stuttering severity.

Table 2. This table is comprised of all individual data measurements collected from each participants' speech samples. The age and gender of each participant is listed as well as their pre-treatment and post-treatment stuttering severity scores (SSI) and naturalness (Nat) ratings. The SFF mean, SFF standard deviation (Std) in hertz (Hz) and in semitones (ST), and interquartile range in hertz and in semitones are also listed for each participant.

ID	Age	Gender	Pre-Tx SSI	Pre-Tx Nat	Post-Tx SSI	Post-Tx Nat	Mean (Hz)	Std (Hz)	Std (ST)	IQR (Hz)	IQR (St)
P01	29	F	28	3.60	15	1.65	200	20.81	1.64	27.13	2.25
P02	24	M	27	6.20	9	5.00	126	17.19	2.21	21.15	2.94
P03	24	F	25	5.70	20	2.86	202	23.79	1.92	33.17	2.85
P04	19	M	34	7.10	10	2.00	158	21.05	2.15	28.69	3.13
P05	23	M	20	2.10	14	1.90	127	13.65	1.75	16.10	2.19
P06	19	M	33	6.00	7	3.05	140	17.49	2.03	23.45	2.90
P07	24	M	35	6.29	6	3.38	107	13.87	2.11	15.98	2.61
P08	33	M	30	4.50	11	3.45	113	20.90	2.92	27.10	4.17
P09	31	M	29	1.20	9	2.15	127	21.39	2.67	26.82	3.68
P10	33	M	25	3.60	0	4.28	108	15.02	2.24	19.56	3.13
P11	21	M	24	5.29	16	6.10	141	13.63	1.58	15.87	1.94
P12	21	M	30	4.50	0	5.35	148	24.25	2.62	28.15	3.28
P13	21	M	22	6.00	9	3.55	145	19.70	2.19	25.57	3.02
P14	24	M	18	2.40	9	2.71	122	12.38	1.66	15.29	2.16
P15	24	M	20	2.20	7	2.47	112	15.05	2.17	17.63	2.72
P16	20	M	23	2.80	6	2.55	140	15.53	1.81	18.81	2.33
P17	38	M	25	4.80	8	6.14	109	18.50	2.69	27.52	4.34
P18	21	M	24	3.30	17	3.23	142	18.80	2.14	24.53	2.98
P19	20	M	27	6.20	17	1.90	112	14.68	2.13	18.11	2.81
P20	21	M	26	5.60	10	2.80	120	11.16	1.53	12.90	1.86
P21	42	M	18	2.10	6	2.29	110	17.61	2.55	21.89	3.43
P22	34	M	29	4.10	8	2.67	111	13.75	2.01	16.64	2.59
P23	28	M	18	1.60	6	2.57	119	27.50	3.58	32.76	4.86
P24	25	M	22	3.20	12	2.00	114	15.36	2.18	16.87	2.57
P25	20	M	23	5.30	6	3.55	102	13.39	2.12	17.71	2.99
P26	29	M	33	7.40	6	3.75	166	17.00	1.68	21.52	2.23
P27	20	M	33	7.10	7	5.29	142	18.22	2.08	21.69	2.65
P28	43	M	20	5.00	4	2.42	193	31.97	2.65	47.64	4.31
P29	19	M	17	2.30	12	1.95	124	12.72	1.68	9.75	1.36

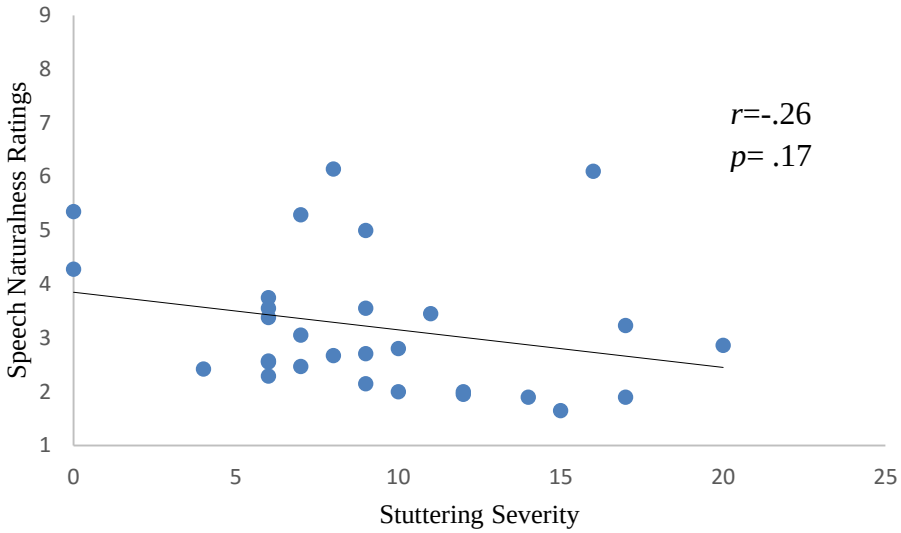

Figure 1. A correlation of r= -.26 and a *p*-value of .17 between stuttering severity and speech naturalness ratings indicates a negative but non-significant relationship.

Relationship Between Behavioral Ratings and SFF Measures

Correlation coefficients were then generated for each of the samples to compare speech naturalness and stuttering severity to each the SFF variation measures. As seen in Table 4, correlation coefficients of SFF variation measures compared to post treatment naturalness were not significant.

Table 4. This table shows the correlation coefficients between pre-treatment and post-treatment stuttering severity ratings and naturalness ratings as well as acoustical measurements of SFF mean, SFF standard deviation in hertz and in semitones, and interquartile range in hertz and in semitones.

	Pre tx SSI	Pre tx Nat	Post tx SSI	Post tx Nat	Mean (Hz)	Std (Hz)	Std (St)	IQR (Hz)	IQR (St)
Pre-Tx SSI	--	.67	-.10	.26	.15	.03	-.07	.06	-.11
Pre-Tx Nat		--	.02	.41	.26	.01	-.23	.12	-.10
Post-Tx SSI			--	-.26	.27	-.14	-.37	-.12	-.32
Post-Tx Nat				--	-.11	-.01	.09	.01	.13

When compared to post treatment naturalness ratings, the correlation for the SFF mean was *r*= -.11 (Figure 1), for the standard deviation the correlation was *r*= -.01, for standard deviation of semitones the correlation was *r*= .09 (Figure 2), for interquartile range the correlation was *r*= .02, and for interquartile range of semitones he correlation was *r*=.13 (Figure 3). None of these relationships proved significant.

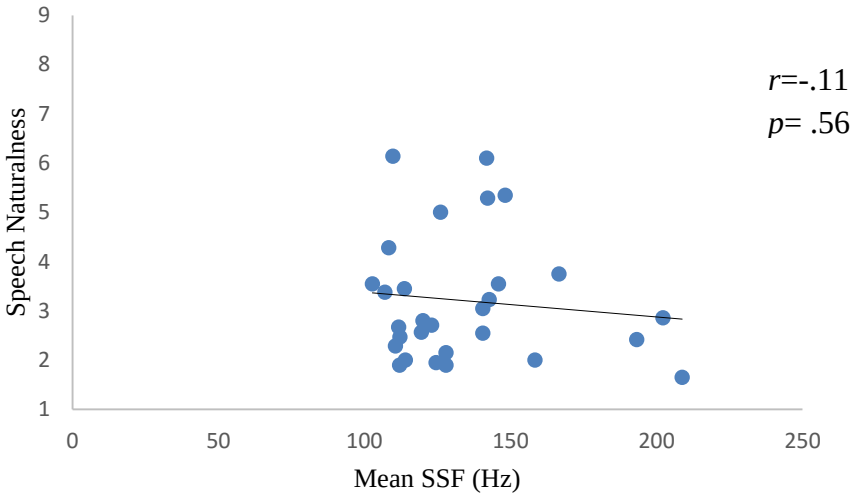

Figure 2. Mean SFF to speech naturalness ratings correlation was -.11 with a P value of .56. Results were not significant.

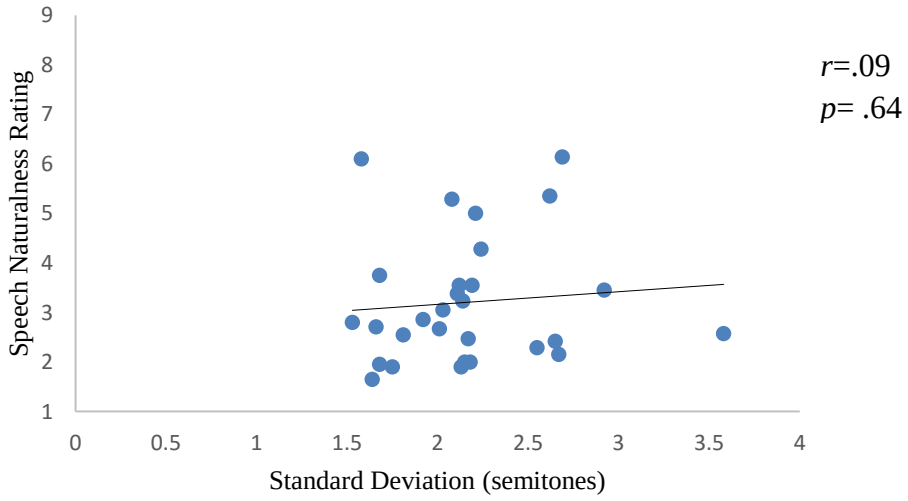

Figure 3. SFF standard deviation in semitones to speech naturalness ratings correlation was *r*=.09 with a P value of .64. Results were not significant.

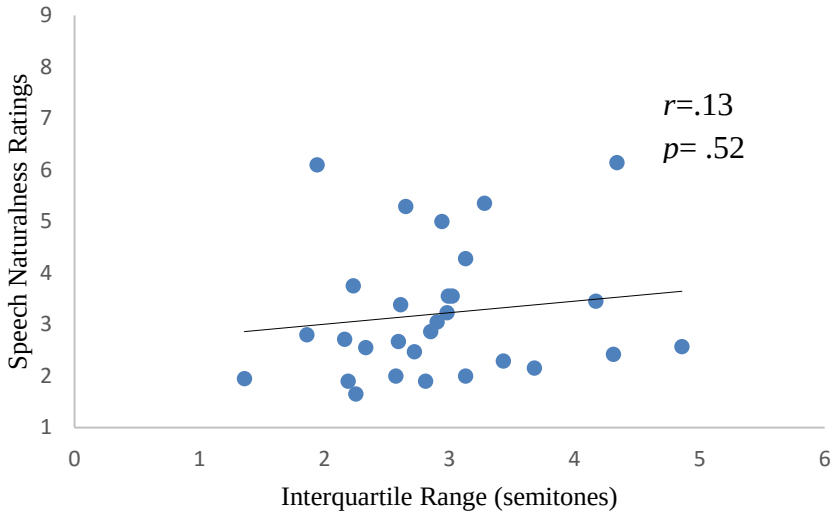

Figure 4. SFF interquartile range in semitones to speech naturalness ratings correlation was *r*=.09 with a P value of .47. Results were not significant.

While there was no relationship observed between SFF variation measurements and ratings of speech naturalness, it is important to note that there was a statistical trend toward a significant negative correlation between SFF variability measures and post-treatment SSI-3 ratings (Figures 4 and 5).

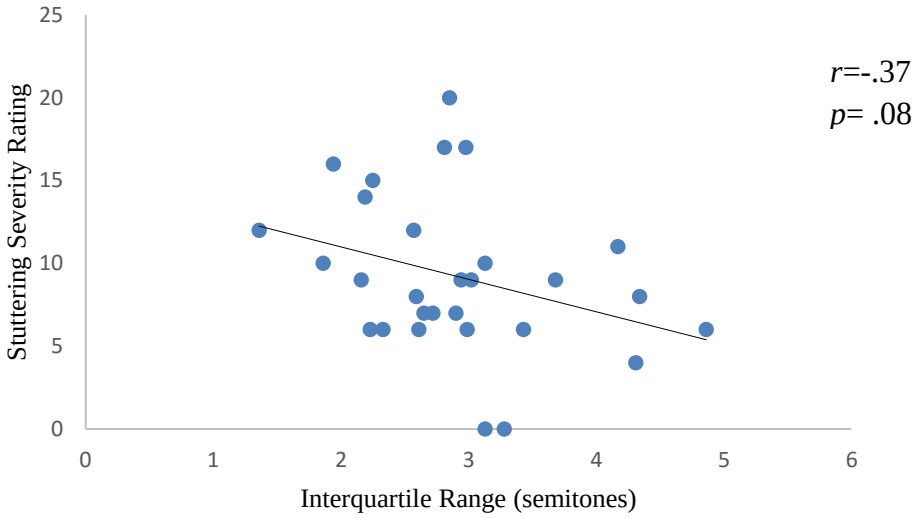

Figure 5. SFF interquartile range of semitones to stuttering severity ratings correlation was *r*=-.37 with a P value of .08. Results were trending toward slight significance.

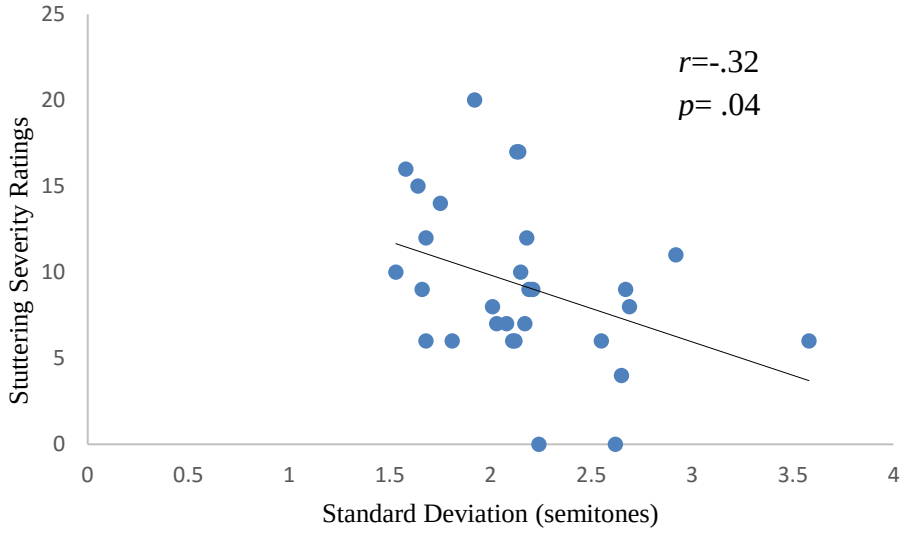

Figure 6. SFF standard deviation in semitones to stuttering severity ratings correlation was *r*=-.32 with a P value of .04. Results were trending toward slight significance.

Discussion

Impact of SFF on Naturalness Ratings

There was no clear relationship found between the SFF variation measures and ratings of speech naturalness. This lack of relationship was contrary to the original hypothesis that predicted smaller variations in SFF would lead to increased ratings of less natural speech.

Although significant relationships did not appear between the SFF variation data and the speech naturalness ratings, a slight negative relationship emerged when comparing post-treatment stuttering severity ratings and the standard deviation of semitones and the interquartile range of the semitones. These relationships imply that individuals with severe stuttering scores had less variation in SFF. One possibility is that those more severely affected participants selected to speak with a more restricted SFF range in an attempt to better control their stuttering.

The comparison of the post-treatment stuttering severity and naturalness ratings indicate that stuttering may not be the key factor in listener's perceptions of speech naturalness. Other factors are clearly at play. The investigation of speaker fundamental frequency influence on speech naturalness does not account for observed changes in listener ratings of speech naturalness, therefore, other acoustic correlates should be investigated.

Limitations

One limitation of the study was the utilization of reading samples for acoustic analysis rather than the monologue samples. The naturalness ratings were based on samples from each participant’s monologue samples. Sometimes adults who stutter experience variations in frequency and type of disfluencies within different contexts of speech. Analyzing portions of the spontaneous speech monologue samples may have yielded different results.

Overall, the analyses may have been influenced by the acoustical characteristics by which fundamental frequency variability was measured. This study focused on broad estimates of SFF variation. It is possible that examining SFF contours at the phrase/utterance level would better capture any anomalies in the speech samples.

Stuttering severity was estimated using the Stuttering Severity Instrument (SSI-3), which provides a composite score based on stuttering frequency, duration of the longest stuttering moments and an estimate of the severity of physical concomitants. The SSI-3 has been criticized due to it psychometric properties. Future studies may consider using other measures associated with stuttering severity such as stuttering frequency.

The sample size (29) may have been more accurately representative of adults who stutter in general with a larger pool of samples. There were also only two females included in the study. This, again, could have impacted the accuracy of the representation of adults who complete intensive fluency shaping programs.

The participants were all military members who had received at least a high school diploma prior to the data collection, but factors related to literacy ability were not accounted for and may have influenced the results. Reader literacy ability may have impacted the fluency of the read aloud passage and potentially impacted the post-treatment naturalness rating.

The study also did not contain data from a control group of typically speaking adults. Speech naturalness ratings and SFF variations from this group would have proven useful to compare with the results of the post-treatment adults who stutter. This comparison is necessary for exploring what type of SFF variations and relationships (if any) between speech naturalness occur in adults who do and do not stutter.

Future Directions

Despite little relationship appearing within the capacities in this analysis, further exploration of SFF impact on speech naturalness could prove interesting if different methods of measurement are pursued. Initial observations of the video samples noted strange pitch and prosodic behaviors that inspired this investigation. This analysis revolved around measurement of variation within each participant's entire reading sample. During the next analysis, deeper focus should target the fluctuation of SFF at the beginning of each utterance, or even each breathe segment.

Future study may benefit from replication of the listener rating procedure with similar populations compared to the original study by Tasko et al (2007) and inclusion of a typically speaking control group. The population of post-treatment adults who stutter is a limiting factor, as it is a difficult population to find and track. Use of a larger participant pool with a variety of speaking fluency could allow for greater likelihood of identifying a relations between acoustic measures and perceived naturalness of speech of post-treatment adults who stutter.

References

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. doi:10.1016/j.jfludis.2008.04.002

Baken, R.J. (1987). *Clinical Measurement of Speech and Voice.* Boston: College-Hill.

Healey, E. (1982). Speaking fundamental frequency characteristics of stutterers and nonstutterers. *Journal of Communication Disorders*, 15(1), 2129. doi:10.1016/0021-9924(82)90041-7

Martin, R. R., Haroldson, S. K., & Triden, K. A. (1984). Stuttering and Speech Naturalness. *Journal of Speech and Hearing Disorders, 49*(1), 53-58. doi:10.1044/jshd.4901.53

Metz, D. E., Schiavetti, N., & Sacco, P. R. (1990). Acoustic and Psychophysical Dimensions of the Perceived Speech Naturalness of Nonstutterers and Post-treatment Stutterers. *Journal of Speech and Hearing Disorders, 55*(3), 516-525. doi:10.1044/jshd.5503.516

Schaeffer, N. (2006). The effects of differential prolongations of phrases on the perceptions of speech naturalness. *Contemporary Issues in Communication Sciences and Disorders*, 33, 152-159.

Schaeffer, N., & Eichorn, N. (2001). The effects of differential vowel prolongations on perceptions of speech naturalness. *Journal of Fluency Disorders, 26*(4), 335-348. doi:10.1016/s0094-730x(01)00101-2

Tasko, S. M., McClean, M. D., & Runyan, C. M. (2007). Speech motor correlates of treatment-related changes in stuttering severity and speech naturalness. *Journal of Communication Disorders, 40*(1), 42-65. doi:10.1016/j.jcomdis.2006.04.002

Yairi, E., & Seery, C. H. (2015). *Stuttering: Foundations and clinical applications* (Second edition). Boston: Pearson.

Appendix 1

Orthographic Transcription of Speech Sample

If you want to grow strong you must exercise. No rules can teach you to use language unless you practice using it. You must develop habits. You can no more learn to read by memorizing rules than you can become a full back by memorizing the rules of football. In both cases, rules will help you to practice, but it is the practice that develops your skill. The exercises in this class are planned to give you the practice. They will help you to form good habits and acquire skills of communication but it will require work on your part. You must practice over and over the skills you want to learn. But even on the easiest ones you must be sure to practice correctly. The exercise can strengthen only what it practices. That stands to reason if you practice a mistake the mistake will grow if you practice a skill the skill will grow.

Total syllables: 208

Appendix 2

F0 Extraction and Editing Procedure

This process is aimed to clean up speaking fundamental frequency data that may be affected by TF-32 algorithm calculation errors. The program currently is attempting to measure f0 at equal intervals for the duration of the sample, but errors can manifest as spikes in the speaking fundamental frequency contour, and if left on the plot, can skew the data.
The following procedure is the editing process for eliminating these spikes in fundamental frequency.

1) Investigator will open a speech sample in the TF-32 program. Go to view >open> pitch trace to display the speaking fundamental frequency plot.
2) Investigator will look for spikes or flares in f0 that will take the form of vertical lines in the contour.
3) If a f0 spike is evident:
    - Enclose the spike with the left and right cursors.
    - Page down to zoom into the selected section.
    - F0 readings are displayed in the upper left hand corner of the plot. To get the frequency for a particular moment on the contour, adjust the left and right cursors. The left cursor will appear as “initial f0”, the right will appear as “final f0.” The difference between the initial and final frequencies will appear as well.
    - While zoomed into the section, the spikes will take the form of a “step like” image or a vertical line amidst a relatively horizontal line. This “step” may be appear as a drastic increase or decrease frequency.
    - Place one cursor on the spike and another on the flat f0 line to compare the f0 initial and f0 final frequencies. If the f0 measurement (upper left corner) differs severely (almost a frequency doubling, i.e. initial=50hz final=100hz, or relatively close), it can be assumed that the sudden change was due to an algorithmic error, and is not a naturally occurring variation in pitch.
4) To eliminate the error:
    - Place right and left cursor around the entire spike or “step.”
    - Right click and hit pitch trace.
    - A box will appear. Hit “zero”.
    - Page up to return to the full sample display.
5) Complete this procedure for all other spikes in f0 in the sample. Focus primarily on eliminating the obvious spikes in f0.