Acoustic Properties of Speech Under Stress in Preschool Children Who Do and Do Not Stutter

Emily Morrow

Vanderbilt University

Abstract

Previous research has shown that stuttering, a potentially life-altering developmental disorder with typical onset during the preschool years, is linked in severity to temperamental and situational emotionality. Thirty-three participants, aged four to six years old, 14 of whom stutter and 19 of whom do not, provided temperamental measures of emotionality via parent-report surveys. Measures of stress/emotionality were derived from acoustic data (fundamental frequency and jitter) drawn during a card stressor task as part of a larger study. Analyses included correlations between temperamental and acoustic measures of emotionality for all participants, as well as comparisons of temperament data and lab acoustic measures of fundamental frequency and jitter between children who do and do not stutter. Although independent samples t-tests and discriminate function analysis showed no significant difference between the two groups for either temperamental or acoustic data, bivariate correlations for both groups showed significant correlations between temperament measures of emotional reactivity and regulation, and acoustic measures, such as mean jitter and jitter range. Results support acoustic measures as indicators of vocal stress in children who do and do not stutter.

Stuttering, a potentially life-altering developmental disorder with typical onset between 2.5 and 6 years of age, involves the disruption of speech patterns through repetitions of sounds, syllables, and single-syllable words, in addition to sound prolongations (e.g. Conture, 2001; Bloodstein & Grossman, 1981; Yairi & Ambrose, 1992). The frequency, severity, and type of these disfluencies can vary both between people and in different speaking situations for the same person (e.g. Bloodstein, 1950; Walden et al, 2012). Stuttering affects about three times as many boys as girls, with 1% of children continuing to stutter after age six and experiencing negative impacts in the academic, social, and vocational spheres of life (e.g. Yairi, 1993). Because of the negative impact of stuttering on day-to-day life, it is important to identify the contribution of social and emotional factors both to the onset of stuttering and to its selective persistence.

Children who stutter (CWS) have been described as more emotionally reactive, less able to regulate their emotions, and less able to regulate their attention than their normally fluent peers (CWNS) (Karrass et al., 2006). Emotional reactivity reflects the tendency to frequently experience high levels of emotional arousal, whereas emotional regulation indicates modulation of the occurrence, intensity, and duration of emotion-based physiological responses to be more congruent or situationally appropriate (Calkins et al., 1999; Thompson, 1994). In addition, stuttering has been linked to speech-language abilities, with CWS being three times more likely than children who do not (CWNS) to exhibit dissociations across speech-language domains (e.g. Anderson & Conture, 2000; Yairi, 1992).

Walden and Conture (2012) developed a dual-diathesis stressor model of developmental stuttering involving emotional and linguistic factors. The model posits that children have individual predispositions on each of two diatheses: the emotional diathesis (reactivity and regulation) and the speech-language diathesis (expressive and receptive). Stressors (emotionally-or linguistically-charged situations) "activate" the emotional and speech diatheses to different degrees depending on each child's specific predisposition. For instance, a public speaking situation–stressful both emotionally and linguistically–would trigger stuttering to a higher degree in a child who is highly emotional and demonstrates diminished speech-language abilities than in a child who shows lower activation levels on the diatheses. In other words, a child's emotional or linguistic characteristics can predict stuttering in stressor situations (Walden et al., 2012).

The empirical study of the dual-diathesis model involved the use of overheard conversations in eliciting emotion, and findings suggest that emotional regulation is especially key in determining the effects of both emotionality and linguistic abilities on stuttering: Regardless of his/her level of parent-reported negative emotionality, a child who demonstrated more regulatory behaviors during the experiment was less likely to stutter. Language measures from the same experiment also demonstrate the importance of regulatory abilities in language development: children with larger variations among scores on the separate language tests administered during the experiment were more likely to stutter. This could result from the inability to regulate attention and emotional arousal in order to concentrate equally and fully on each task (Walden et al., 2012). As these results suggest that emotionality may in fact underlie both the emotional diathesis and the language diathesis as a causal factor, a focus on the interaction between emotional reactivity and regulation in both talker groups (CWS and CWNS) becomes especially important in identifying major contributing factors to childhood stuttering.

Emotional reactivity and regulation have been measured by self- or other-report (as an indicator of overall emotional temperament), behavioral observations and physiological data. In studies involving young children, subscales of the Children's Behavior Questionnaire (CBQ) and the Behavioral Style Questionnaire (BSQ), completed by parents about their children, have been established to form additive composite scales to measure temperamental levels of both emotional reactivity and emotional regulatory abilities (BSQ; CBQ; Rothbart et al., 2001). Such measures include emotional reactivity, emotional regulation, and attention regulation from the BSQ and negative affectivity and effortful control (the ability to inhibit a dominant emotional response to make it more situationally congruent) from the CBQ (Rothbart et al., 2001). Scores on these composite measures indicate each participant's overall emotional temperament. Results of previous parent-report research have shown differences in emotionality between talker groups and a significant interaction between parent-reported reactivity and regulation on stuttering. Karrass et al. (2006) found that CWS demonstrated higher levels of emotional reactivity and lower levels of emotional regulation using parent report on the BSQ. In the Walden and Conture (2012) study, there was an interaction between negative emotion and regulation on stuttering: When high levels of negative emotionality and emotional regulation (as reported by parents on the CBQ) co-occurred, all children (CWS and CWNS) demonstrated less stuttering.

Although survey measures of overall temperament are useful in determining emotionality, physiological data collected during laboratory tasks can also prove helpful in providing an emotional profile, both by determining each individual's response level to specific situations and by comparing task-specific reactivity across participants. Because stuttering is a psycholinguistic disorder, the use of vocal acoustic analyses of stress could prove especially informative in examining its emotional causation. Acoustic analysis identifies specific components of emotion in speech, describing the characteristics of the speech wave (Juslin & Scherer, 2008). Though speech is understood as the combined physiological effort of three subsystems–respiratory (i.e. lungs), phonatory (i.e. larynx), and articulatory (i.e. tongue)–somewhat-recent technological advances (specifically the digitization of speech sounds) have made analyses of these complex acoustic properties of speech possible, and the field has grown significantly in the last 50 years (Kent & Read, 1992).

In any such vocal analysis, it is important to note that emotion is in fact a process of mental and physiological events, not a fixed, unified condition (Scherer, 1984). Scherer defines emotion as “the organism’s interface with the world outside” and identifies three main factors: evaluation of the relevance and significance of a stimulus, physiological and psychological preparation to take a certain action (defined as emotional “activity” in most literature), and communication of emotional state to surrounding organisms (Scherer, 1981, Murray & Arnott, 1992). Many acoustic analyses focus on component models of emotion, which allow for distinct analysis of emotion via vocal acoustic analysis. Acoustic cues serve as an intermediary step between the expression of an emotion by the speaker and its interpretation by the perceiver in Scherer’s component process theory of emotional conveyance (Scherer, 1986).

The assumption is that emotional arousal causes physiological changes in respiration, phonation, and articulation that produce emotion-specific patterns of acoustic parameters (Scherer, 1986). Thus, emotional activity dictates the qualities of vocal communication: Higher levels of emotional arousal will cause the larynx to contract, leading to changes in specific vocal parameters, especially fundamental frequency and jitter (Kent & Read, 1992). Consequently, levels of these parameters provide measures of emotional reactivity via activation of the important acoustic measures in determining vocal stress. As the field has grown, such vocal analyses have identified specific aspects of this physiological effort that may indicate an emotional basis for certain speech utterances, turning speech analysis into a useful tool for measuring emotionality (Juslin & Scherer, 2008, Scherer, 2003). In emotional acoustics, vowels generally comprise the peak volume intensity of speech (as measured in decibels), and are thus especially useful for analysis because vocal nuances are most perceptible in highly voiced sounds (Kent & Read, 1992). Researchers in the field have established a model of studying emotionality in voice via specific acoustic parameters, including fundamental frequency and jitter.

Fundamental frequency (f0), the number of cycles per second in a periodic sound, is perceived by the human ear as pitch and has been shown to rise with increased levels of emotion (e.g. Bachorowski & Owren, 1995; Kent & Read, 1992; Scherer, 1986; Bachorowski et al., 2001; Scherer, 2003). This rise results from the contraction of the larynx with increased emotional arousal (Kent & Read, 1992). f0 is the most established and widely-researched measure of vocal emotionality, with base and mean fundamental frequency demonstrated to rise with vocal stress (Scherer, 2003). Research in the area often involves the comparison of ECG (electrocardiogram, or heart rate) levels of emotional arousal with changes in vocal quality. Johnstone and Scherer (1999) found a mean increase in fundamental frequency floor and fundamental frequency mean with higher levels of emotional arousal when participants were asked to imagine themselves in specific emotional states (e.g. both ECG and F0 showed significantly higher levels for anger than for boredom). They found their results to be consistent with previous findings that contraction of the larynx with higher emotional arousal leads to less glottal airflow and more high-pitched speech.

Jitter (cycle-to-cycle perturbation in f0) has also has been shown to be affected by emotion, and high levels of jitter have been found to correspond with higher levels of f0 (e.g. Scherer, 1986). In their 1999 emotional imagination study, Johnstone and Scherer found that the perturbations in f0 (or jitter) tended to rise with higher mean levels of F0.

However, despite the general rise of fundamental frequency and jitter with increased levels of emotional arousal, a special exception for stress/anxiety has been consistent in research using acoustic techniques to analyze different emotions. In their study, Johnstone and Scherer (1999) found that though fundamental frequency floor consistently rose across levels of increased emotional arousal, f0 range and jitter actually decreased, rather than increased, with higher levels of anxiety. Their findings were consistent with previous research using acoustic analysis to measure these specific signs of vocal stress (e.g. Smith, 1977).

Despite the properties of acoustic analysis that make it a very specific tool to emotionality stuttering research, limitations of the analysis process necessitate a specific type of data in order to use it. Specifically, direct comparisons can only be made between participants uttering the same vowel sound in the same word in the same placement within the utterance; tasks involving variable or participant-generated speech do not allow such direct comparison (Murray & Arnott, 1992; Juslin & Scherer, 2008). In addition, moving beyond a single measure of acoustic properties is beneficial in ensuring a valid measure of the acoustic parameters (Juslin & Scherer, 2008). Thus, acoustic analysis is an effective tool for comparison when researchers dictate participants' speech patterns. However, emotional responses are most valid when naturally produced by participants; thus, laboratory manipulation of emotional state could allow for the controlled recording of acoustic responses, though little research of this type exists in the field to date (Murray & Arnott, 1992, Juslin & Scherer, 2008). To date, natural vocal states achieved through laboratory manipulation have generally been consistent with, if less significant than, parameters achieved through portrayal by professional actors (Juslin & Scherer, 2005). Using acoustic analysis of participant vocal responses to a laboratory video game, Johnstone et al. (2005) found acoustic changes consistent with previous laboratory research using actor portrayals: Participants' f0 tended to rise in higher-importance or less-congruent situations in a game, suggesting that laboratory manipulations of emotion could be promising in introducing natural emotion to the field of acoustic analysis. Because of these limitations on the use of acoustic analysis, previous research involving acoustic analysis for children who stutter is slim, and much of it focuses on prosody and duration of their speech utterances rather than vocal parameters (Zebrowski et al., 1985). Thus, acoustic analysis provides an innovative measure of emotionality for the present study, while also allowing further development of a new genre of stuttering research.

The goal of the present study is to further examine emotionality in childhood stuttering by using acoustic measures of stress/arousal in a standard laboratory stress task and comparing temperamental and acoustic measures of emotionality between children who do and do not stutter. By combining acoustic measures of emotionality with temperamental measures of emotionality, the study examines the validity of both by considering the relation between them, while also examining the relation between temperamental emotionality and emotion under stress in children who do and do not stutter. Thus, the correlation between the two types of measures will be especially important here, as shown in the first hypothesis:

1. Temperamental and acoustic measures of emotion should be correlated, both for children who do stutter and for children who do not.

In addition, as children who stutter have demonstrated higher levels of emotional reactivity and lower levels of emotional regulation in previous research, the same should be true for both types of measures here:

2. Children who stutter should demonstrate higher levels of temperamental emotional reactivity and lower levels of temperamental emotional regulation than children who do not stutter.
    a. On the CBQ, CWS should have higher scores of negative affect and lower scores of effortful control than CWNS.
    b. On the BSQ, CWS should have higher scores of emotional reactivity and lower scores of emotional regulation and attention regulation than CWNS.
3. Children who stutter should demonstrate higher levels of emotionality than children who do not stutter, as measured by acoustic measures of speech during a laboratory stressor task.
    a. CWS should have higher overall mean fundamental frequency than CWNS.
    b. CWS should have lower variability in fundamental frequency (fundamental frequency range) than CWNS.
    c. Mean jitter and jitter range should differ between CWS and CWNS.

Method

Participants

Participants were 33 children aged 4 to 6 years selected from the Vanderbilt Developmental Stuttering Project, a larger longitudinal study. Participants were selected based on parent survey responses and words named during a laboratory card-naming stressor task. Ninety-eight participants were considered for inclusion at their first visit. Twenty-seven participants who did not name the chosen words in the proper order were excluded from the sample. In addition, 22 participants without a useable physiological file for acoustic data from the card stressor task were excluded from the study, and 22 participants were removed for lack of temperament data. Eleven of the selected participants at the first time point stuttered, as demonstrated by their percentage of stuttered utterances (during a conversation with a clinician) and their scores on the Stuttering Severity Instrument (SSI-4) at the chosen time point (8 male, 3 female), and 16 did not (12 male, 4 female). CWS had scores above 10 (st least "mild") on the SSI-4 or more than 3% stuttered utterances during the conversation with the clinician. Fifteen children who did not name the correct words at time point 1, but who had later attended time point 2, were considered for inclusion at time point 2. Of those considered, 8 lacked physiological data, 1 lacked temperament data, and 6 were selected for inclusion: 3 CWS (all male) and 3 CWNS (2 male, 1 female).

Measures

Temperament

Temperamental emotionality was measured via parent responses to two questionnaires: the Behavioral Style Questionnaire (BSQ) and the Child Behavior Questionnaire (CBQ). Parents completed the questionnaires, either on paper or online, when their children completed the laboratory card-naming stressor task. Surveys were scored using SPSS syntaxes developed to extract measures of emotionality: emotional reactivity, emotional regulation, and attention regulation from the BSQ; negative affect and effortful control from the CBQ. Two versions of the CBQ (one with 70 questions and one with 243 questions) were administered to different families at the first time point, but syntax was created to derive the measures from either version. Scores produced by the two versions were highly correlated when the syntax were run on the same data (r = .773).

Acoustic Measures of Stress/arousal

Measures of emotion/stress in the voice were obtained from acoustic files extracted from a laboratory card-naming task completed at the first time point of the longitudinal study. Words selected for analysis had “long a” vowels consisting of “grapes,” “cupcake,” “airplane, and “cake.” Of the words available for the list, experimenters ranked preference of all possible words to determine which three of the available word options were used for each participant’s list of 3 words. In ranking the words, experimenters gave preference to words spaced more evenly throughout the task in order to gather acoustic data at the beginning, middle and end of the task. Acoustic data was captured with AcqKnowledge, cropped to include only selected words using Audacity, and analyzed using a script on the acoustic software PRAAT. Experimenters manually cropped acoustic data on Audacity and isolated selected vowels for analysis on PRAAT; reliability was established between two acoustic coders. The PRAAT script created for the experiment produced measures of fundamental frequency, jitter, and duration from the manually selected vowel segments (Appendix A).

Design

Talker group (CWS/CWNS) is the independent variable. Dependent variables include acoustic measures of fundamental frequency and jitter (mean and range), as well as temperamental variables from the BSQ and CBQ, including emotional reactivity, emotional regulation, and attention regulation for the BSQ and negative affect and effortful control from the CBQ**.** Correlations among the acoustic measures and temperament variables were conducted both across the two groups and within each group.

Procedure

Temperament

Parents completed the two surveys about their children, either on paper or online, either at the data-collection visit or within a few days of the visit when requested by email. The survey scores used for analysis came from the same time point as when the children completed the card stressor task.

Stress Task

During the task, the experimenter, who had already developed a rapport with the child, presented a series of thirty picture cards selected at varying levels of difficulty from the Peabody Picture Vocabulary Test (PPVT-IV) and asked the child to name each picture. The cards were always presented in the same order. The child spoke into a microphone to record acoustic data but left hands steady so as not to disturb electrodes recording physiological data. The entire task lasted about a minute. The experimenter quickly and loudly slapped each card onto the table and encouraged the child to move quickly throughout the task, using phrases like “Go faster!” and “You can go faster than that!” at regular intervals, in order to create social and time pressure. Many experimenters also presented the task as a race by asking the child if he or she could be the quickest ever to complete the task.

Results

Data Analyses

Correlations for hypothesis 1 were two-tailed bivariate correlations between acoustic and temperamental variables for both CWS and CWNS. Because histograms of some acoustic measures, particularly jitter and jitter range, showed a Poisson distribution, the Spearman correlation was used.

Hypotheses 2 and 3 were tested with independent samples t-tests between the two groups for both acoustic and temperamental variables.

Descriptive Statistics

Overall and group means for CWS and CWNS are in Tables 1-3. No significant differences were found between the two groups, for temperamental measures (i.e., BSQ emotional reactivity, BSQ emotional regulation, BSQ attention regulation, CBQ negative affect, CBQ effortful control), acoustic measures (i.e., mean and range of fundamental frequency, mean and range of jitter), or for descriptive measures (i.e., gender, age, percent SLD's per 100 words, SSI score).

Bivariate Correlations

No difference in correlations between CWS and CWNS were found using the Fisher r-to-z transformation, so correlations reported here reflect both groups combined.

Table 4 shows the correlations among the temperamental variables (i.e., BSQ and CBQ scores) for both groups combined. Correlations were significant between several temperamental variables, including BSQ Emotional Regulation and BSQ Emotional Reactivity (Spearman's rho=-.645), BSQ Emotional Regulation and BSQ Attention Regulation (Spearman's rho=.428), BSQ Emotional Regulation and CBQ Effortful Control (Spearman's rho=.454), BSQ Emotional Regulation and CBQ Negative Affect (Spearman's rho=-.725), BSQ Emotional Reactivity and CBQ Negative Affect (Spearman's rho=.654), BSQ Attention Regulation and CBQ Effortful Control (Spearman's rho=.434), and BSQ Attention Regulation and CBQ Negative Affect (Spearman's rho=-.415). The large number of significant correlations among the temperamental variables indicates that they are highly related to each other, a colinearity which may influence interpretation of the results.

Table 5 shows correlations between temperament and acoustic variables (i.e., mean fundamental frequency, range of fundamental frequency, mean jitter, and range of jitter). Several correlations were significant, including mean jitter and BSQ emotional reactivity (Spearman's rho=.468), mean jitter and BSQ emotional regulation (Spearman's rho=-.551), mean jitter and BSQ attention regulation (Spearman's rho=-.465), mean jitter and CBQ negative affect (Spearman's rho=.566), range of jitter and BSQ emotional reactivity (Spearman's rho=.463), range of jitter and BSQ emotional regulation (Spearman's rho=-.469), and range of jitter and CBQ negative affect (Spearman's rho=.462). Though there were no significant correlations involving fundamental frequency, hypothesis 1 is partly supported by the correlations between jitter variables and temperamental variables.

Mean fundamental frequency and age were negative correlated (Spearman's rho=-.361). In addition, range of jitter and mean fundamental frequency were positively correlated (Spearman's rho=.701).

T-Tests and Discriminant Function Analysis

A discriminate function analysis was also used to determine which variables might distinguish between the two groups, but was found to be nonsignificant.

Independent samples t-tests and discriminant function analysis between the two groups involving descriptives (i.e., talker group, gender, age, SLD's per 100 words, SSI score), acoustic variables (i.e., mean fundamental frequency, range of fundamental frequency, mean jitter, and range of jitter), and temperament variables (i.e., BSQ emotional reactivity, BSQ emotional regulation, BSQ attention regulation, CBQ negative affect, and CBQ effortful control) were not significant. As such, hypotheses 2 and 3 are not supported by the results.

Discussion

The present study extends previous research on emotional reactivity, emotional regulation, and stuttering by examining emotion through both temperamental and laboratory stressor measures. Further, findings support the validity of laboratory acoustic measures of stress in children who stutter.

Temperamental and Acoustic Measures

Two acoustic measures, mean jitter and jitter range, measured during a laboratory stressor task were correlated with several parent-report temperament measures, including BSQ emotional reactivity, BSQ emotional regulation, and CBQ negative affect. These findings support hypothesis 1 (which was a non-directional hypothesis due to a lack of definitive previous research on jitter and emotion), and also suggest a directionality of emotion's influence on jitter. Particularly, mean jitter and jitter range were positively correlated with measures of reactivity (i.e., BSQ emotional reactivity, CBQ negative affect) and negatively correlated with measures of regulation (i.e., BSQ emotional regulation). Both baseline jitter and range of jitter rose with increased levels of temperamental emotional reactivity and decreased levels of temperamental emotional regulation.

Whereas previous research suggested that mean jitter rises with mean fundamental frequency (e.g. Scherer, 1986), present findings instead support a positive correlation between jitter range and mean fundamental frequency, though there was no correlation between mean jitter and mean fundamental frequency. Jitter itself is a measure of cycle-to-cycle perturbation in pitch, so previous findings have indicated that cycle-to-cycle variability rises with higher baselines of pitch. The present findings also show a positive correlation between mean fundamental frequency and cyclical fundamental frequency variability, but range of jitter instead represents the variability in cyclical changes in fundamental frequency throughout the card naming task, as opposed to the mean cyclical changes during the task as a whole. Consequently, the correlation between mean fundamental frequency and range of fundamental frequency suggests that the amount of cyclical variation in fundamental frequency can differ between individuals, and that changes in that amount of variation throughout the card naming task might be linked to levels of emotional reactivity. Children who are more reactive demonstrate not only more cyclical acoustic change, but also more variation in the amount of cyclical acoustic change throughout the task.

Further, the positive correlation between emotional reactivity and jitter range under stress contrasts with previous research suggesting that the perturbations in pitch often decrease with increased levels of stress, though they would rise with increased levels of other emotions, such as fear and anger (e.g., Johnstone and Scherer, 1999). Our findings suggest that, in contrast with previous research suggesting that stress follows a unique pattern of increased stability of acoustic measures with higher emotionality, jitter range follows the typical pattern of increased emotional arousal by rising with increased emotional reactivity, indicating more vocal perturbations with a stronger reaction to emotional stress. However, because range of jitter and mean fundamental frequency were positively correlated in this study, the positive correlation between range of jitter and emotional reactivity does indirectly support previous findings suggesting that mean fundamental frequency rises with increased levels of reactivity or emotional stress (e.g., Johnstone and Scherer, 1999). Thus, the results of the present study indicate that stress might not be unique in its patterns of reactivity, and might follow the typical pattern of increased measures of arousal with higher levels of stress arousal. Further study is certainly needed to clarify the role stress plays in changing levels of vocal parameters.

These findings generally support mean jitter and range of jitter as laboratory acoustic measures of temperamental emotionality while also suggesting a directionality the relation between emotional reactivity, emotional regulation, and perturbations in jitter. Future study is certainly necessary both to further examine and establish this directionality and to examine the relation between fundamental frequency measures and stress.

Limitations

As part of a larger longitudinal study, this study was limited by the number of participants with necessary data to be included in the analysis. Future research should include more participants for more power to detect relations between acoustic measures of stress and temperamental levels of emotionality. This might further establish jitter as a measure of temperament while also supporting previous research on the validity of fundamental frequency as a measure of temperament.

Further, acoustic data can be volatile, and future research should control for factors that might alter results, especially the distance of the child's mouth from the microphone. Varying distances from the microphone produce varying vocal intensities, which could influence other acoustic measures. Gathering a baseline intensity by having the participant read a list of words in a non-agitated state before and after the task could prove helpful in eliminating this concern.

Additionally, gathering a baseline fundamental frequency from the same vocal sample would control for individual differences in fundamental frequency. As noted above, mean fundamental frequency and age were correlated in the sample (Spearman's rho=-.361). Though t-tests revealed no significant difference in age between the groups, the relation between age and fundamental frequency might be a factor in the higher fundamental frequencies of some participants. Having children repeat the exact same word both during the baseline collection and at many points during the task would allow for easier direct comparison and strengthen results.

Concluding Remarks and Future Directions

The present study was motivated by previous research regarding emotionality in combination with both stuttering and acoustic measures of stress. By combining these earlier bodies of work, the study assesses the possibility of using acoustic analysis to measure vocal stress in children who do and do not stutter in the future. Present results support this method, but results should be taken cautiously until replicated in a larger and more controlled study.

Importantly, the findings of this study suggest that both mean jitter and variations in jitter tend to rise with temperamental emotional reactivity and fall with temperamental emotional regulation, though it should be noted that the BSQ measures of emotional reactivity and regulation are highly negatively correlated. This suggests that either the two might reflect opposite end of a single continuum, or that increased regulatory skill substantially dampens emotional reactivity. These results suggest not only a directionality of jitter response to emotional stress, but also that temperamental levels of emotionality truly manifest themselves in stressful situations, both for children who stutter and children who do not. Present findings are consistent with a model of stuttering involving linguistic and emotional diatheses in responding to stressful situations (e.g., Walden et al., 2012). Because the laboratory task was stressful both emotionally and linguistically, the response of these acoustic measures of stress supports the presence of an emotional diathesis in the response of both children who stutter and children who do not to situational stress.

In sum, present findings not only support acoustic analysis as a measure of emotional stress in children who do and do not stutter, but also provide further empirical evidence for the presence of an emotional diathesis in response to situational stress. Future research is necessary to further support these findings, but this initial inquiry provides support for the combination of emotional, acoustic, and stuttering research.

References

Acqknowledge Data Acquisition and Analysis Software. (Version 4.1). [Software]. BIOPAC Systems.

Anderson, J. D., & Conture, E. G. (2000). Language abilities of children who stutter: A preliminary study. *Journal of Fluency Disorders*, *25*, 283–304.

Audacity: Free audio editor and recorder. Retrieved June 9, 2013, from audacity.sourceforge.net.

Bachorowski, J.A., & Owren, M.J. (1995). Vocal expression of emotion: acoustic properties of speech are associated with emotional intensity and context. *Psychological Science, 6,* 219-224.

Bachorowski, J.A., Somski, M.J., & Owren, M.J. (2001). The acoustic features of human laughter. *Journal of the Acoustic Society of America, 110,* 1581-1597.

Bloodstein, O., & Grossman, M. (1981). Early stuttering: Some aspects of their form and distribution. *Journal of Speech and Hearing Research*, *24*, 298–302.

Bloodstein, O., & Bernstein-Ratner, N. (2008). A handbook on stuttering (6th ed.). Clifton Park: Delmar.

Bloodstein, O. (1950). A rating scale study of conditions under which stuttering is reduced or absent. *Journal of Speech and Hearing Disorders, 15,* 29-36.

Boersma, Paul (2001). PRAAT, a system for doing phonetics by computer. *Glot International* 5:9/10, 341-345.

Calkins, S.D., Gill, K.L., Johnson, M.C., & Smith, C.L. (1999). Emotional reactivity and emotional regulation strategies as predictors of social behavior with peers during toddlerhood. *Social Development, 8,* 310-334.

Conture, E. G. (2001). Stuttering: Its nature, diagnosis, and treatment. Boston: Allyn & Bacon.

Dunn, L., & Dunn, L. (1997). Peabody Picture Vocabulary Test-III (PPVT-III) (3rd ed.). Circle Pines: American Guidance Service, Inc.

Fairbanks, G., & Pronovost, W. (1939). An experimental study of the pitch characteristics of the voice during the expression of emotion. *Speech Monograph, 6:* 87-104.

Johnstone, T., & Scherer, K. R. (1999). The effects of emotions on voice quality. In *Proceedings of the XIVth International Congress of Phonetic Sciences* (pp. 2029-2032). San Francisco: University of California, Berkeley.

Johnstone, T., van Reekum, C., Hird, K., Kirshner, K., & Scherer, K. (2005). Affective speech elicited with a computer game. *Emotion (5):* 513-518.

Juslin, P.N., & Scherer, K.R. Speech emotion analysis (2008). *Scholarpedia, 3(10): 4240.*

Karrass, J., Walden, T.A., Conture, E.G., Graham, C. G., Arnold, H.S., Hartfield, K.N., & Schwenk, K.A. (2006). Relation of emotional reactivity and regulation to childhood stuttering. *Journal of Communication Disorders, 39,* 402-423.

Kent, R.D., & Read, C. The acoustic analysis of speech. San DiegoL Singular/Thompson Learning.

McDevitt SC, Carey WB. (1978). The measurement of temperament in 3 7-year-old children. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 19*, 245–253.

Murray, I.R., & Arnott, J.L. (1992). Toward the simulation of emotion in synthetic speech: a review of the literature on human vocal emotion. *Journal of the Acoustic Society of*

America, 93, 1097-1108.

Riley, G. (1994). Stuttering severity instrument for young children-3 (3rd ed.). Austin: Pro-Ed.

Rothbart, M. K., Ahahdi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: the children's behavior questionnaire. *Child Development, 72*,1394–1408.

Rothkrantz, L.M., Wiggers, P., van Wees, J.A., & van Vark, R. (2004). Voice stress analysis. *Text, speech, and dialogue: Lecture notes in computer science, 3206,* 449-456.

Scherer, K.R. (1984). Emotion as a multicomponent process: a model and some cross-cultural data. *Social psychology, 5,* 37-63.

Scherer, K.R. (1981). Speech and emotional states. In Darby, J.K. (Ed.), *Emotions in Personality and Psychopathology*. New York: Grune and Stratton.

Sherer, K.R. (1986). Vocal affect expression: a review and a model for future research. *Psychological bulletin, 99,* 143-165.

Scherer, K.R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, *40,* 227-256.

Smith, G.A. (1977). Voice analysis for the measurement of anxiety. *British Journal of Medical Psychology, 50:* 367-373.

Thompson, R. A. (1994). Emotion regulation: A theme in search of definition. In N. A. Fox (Ed.), The development of emotion regulation: Biological and behavioral considerations. *Monographs of the Society for research in child Development, 59* (2-3, Serial No. 240), 25-52.

Walden, T.A., Frankel, C.B., Buhr, A.P., Johnson, K.N., Conture, E.G., & Karrass, J.M. (2012). Dual diathesis-stressor model of emotional and linguistic contributions to developmental stuttering. *Journal of Abnormal Child Psychology, 40,* 633-644.

Yairi, E., & Ambrose, N. (1992). Onset of stuttering in preschool children: selected factors. *Journal of Speech and Hearing Research, 35,* 782-788.

Yairi, E. (1993). Epidemiological and other considerations in treatment efficacy research with preschool age children who stutter. *Journal of Fluency Disorders, 18,* 197-219.

Zebrowski, P.A., Conture, E.G., & Cudahay, E.A. (1985). Acoustic analysis of young stutterers' fluency: Preliminary observations. *Journal of Fluency Disorders, 10,* 173-192.

Appendix A

PRAAT Script for Acoustic Analysis

```
clearinfo
form Calculate_Acoustic Measures
        #indicate where your sound files and TextGrid are
        sentence input_folder /[insert input folder here]
        #indicate where you want your output to be saved
        sentence output_folder /[insert input folder here]
        #indicate the number of the tier where phrases were annotated
        integer tier 1
        #indicate the symbol used for pause
        word pause
endform
myList = Create Strings as file list... liste 'input_folder$'/*.wav
ns = Get number of strings
line$="FILE'tab$'vowel'tab$'F0'tab$'sdPitch'tab$'jitter'tab$'shimmer'tab$'Intensity'tab$'duration'
newline$'"
line$>'output_folder$'/output data.txt
for i from 1 to ns
clearinfo
        select Strings liste
        name$ = Get string... 'i'
        Read from file... 'input_folder$'/'name$'
```

```
	mySound=selected("Sound")
	mySound$=selected$("Sound")
	nameraw$ = name$ - ".wav"
	nametxg$ = nameraw$ + ".TextGrid"
	Read from file... 'input_folder$'/'nametxg$'
	myTextGrid=selected("TextGrid")
	myTextGridname$=selected$("TextGrid")
#sound = Read from file... 'soundFileName$'
	select myTextGrid
	nInt = Get number of intervals... tier

	select mySound
	To Intensity... 100 0
n=0
for int from 1 to nInt
	select myTextGrid
	int$=Get label of interval... tier int
if int$!= pause$
		select myTextGrid
		start = Get starting point... tier int
		end = Get end point... tier int
#AutoCorrelation=optimized for intonation analysis (pitch etc), CrossCorrelation=optimized for voice analysis (jitter, shimmer etc).
```

```
#Pitch Range Settings: The default settings in Praat are 75-500 Hz. For a male, a reasonable range is 75-300 Hz, for a female, 100-500 Hz. For children ages 4-10: 100-600 Hz (citation: Quantitative Analysis of Pitch in Speech of Children with Neurodevelopmental Disorders. http://www.cs.rochester.edu/u/emilypx/Interspeech12-GK.pdf)
#These are just estimates, you can determine the pitch range by playing with pitch settings until you get the pitch line halfway up the window.
#INTONATION MEASURES (PITCH), WE USE AUTO-CORRELATION FOR PITCH
#100 and 600 HZ are the pitch range settings.
select mySound
pitch1 = To Pitch... 0.01 100 600

meanPitch = Get mean... start end hertz
sdPitch = Get standard deviation... start end hertz
#VOICE MEASURES (JITTER, SHIMMER), WE USE CROSS-CORRELATION FOR JITTER AND SHIMMER
#100 and 600 HZ are the pitch range settings.
select mySound
pitch2 = To Pitch (cc)... 0.01 100 15 no 0.03 0.45 0.01 0.35 0.14 600
plus mySound
pulses = To PointProcess (cc)
plus mySound
plus pitch2
voiceReport$ = Voice report... start end 100 600 1.3 1.6 0.03 0.45
```

```
report$ = Voice report... start end 100 600 1.3 1.6 0.03 0.45
jitter_loc = extractNumber (report$, "Jitter (local): ") *100
shimmer_loc = extractNumber (report$, "Shimmer (local): ") *100
#INTENSITY
select Intensity 'mySound$'
meanIntensity = Get mean... start end dB
#DURATION
dur = end-start
line$="'mySound$"tab$"int$"tab$"meanPitch:3"tab$"sdPitch:3"tab$"jitter_loc:3"tab$"shimmer_loc:3"tab$"meanIntensity:3"tab$"dur:4"tab$"newline$"'
line$>>'output_folder$'/output data.txt
endif
endfor
endfor
```

Tables and Figures

Table 1

Means and Standard Deviations for CWS and CWNS Combined

Variable	N	Mean	Std. Deviation
Age	33	55.8939	6.42266
SLD's per 100 Words	33	4.4945	4.48813
SSI Score	33	11.88	6.918
Mean Fundamental Frequency	33	299.0648	78.06861
Range of Fundamental Frequency	33	108.6988	78.26346
Mean Jitter	33	1.4063	1.38583
Range of Jitter	33	1.7354	3.54913
BSQ Emotional Reactivity	33	3.8067	.74053
BSQ Attention Regulation	33	3.7103	.67121
BSQ Emotion Regulation	33	3.9309	.65035
CBQ Rothbart Effortful Control	33	4.5761	.72542
CBQ Rothbart Negative Affect	33	3.4385	.92483
CBQ Walden Negative Affect	33	3.5424	1.01158
CBQ Walden Emotional Regulation	33	4.7697	.81647

Table 2

Means and Standard Deviations For CWNS Only

Variable	N	Mean	Std. Deviation
Age	19	56.8947	7.27075
SLD's per 100 Words	19	1.7711	2.66395
SSI Score	19	6.84	3.005
Mean Fundamental Frequency	19	298.9448	91.39699
Range of Fundamental Frequency	19	115.7348	66.05209
Mean Jitter	19	1.4486	1.63343
Range of Jitter	19	1.7039	4.26636
BSQ Emotional Reactivity	19	3.7168	.70922
BSQ Attention Regulation	19	3.7537	.70405
BSQ Emotion Regulation	19	4.0616	.70741
CBQ Rothbart Effortful Control	19	4.6326	.86278
CBQ Rothbart Negative Affect	19	3.3584	.97999
CBQ Walden Negative Affect	19	3.4521	1.06692
CBQ Walden Emotional Regulation	19	4.8389	.91469

Table 3

Means and Standard Deviations For CWS Only

Variable	N	Mean	Std. Deviation
Age	14	54.5357	4.98931
SLD's per 100 Words	14	8.1907	3.76883
SSI Score	14	18.71	4.232
Mean Fundamental Frequency	14	299.2278	58.61781
Range of Fundamental Frequency	14	99.1500	94.15883
Mean Jitter	14	1.3489	1.01343
Range of Jitter	14	1.7782	2.40842
BSQ Emotional Reactivity	14	3.9286	.79097
BSQ Attention Regulation	14	3.6514	.64511
BSQ Emotion Regulation	14	3.7536	.53795
CBQ Rothbart Effortful Control	14	4.4993	.50362
CBQ Rothbart Negative Affect	14	3.5471	.86808
CBQ Walden Negative Affect	14	3.6650	.95636
CBQ Walden Emotional Regulation	14	4.6757	.68260

Table 4

Temperament x Temperament Correlations			BSQ Attention Regulation	BSQ Emotion Regulation	CBQ Rothbart Effortful Control	CBQ Rothbart Negative Affect
Spearman's rho	BSQ Emotional Reactivity	Correlation	-.226	**-.645****	-.207	**.654****
		Sig. (2-tailed)	.206	.000	.248	.000
		N	33	33	33	33
	BSQ Attention Regulation	Correlation	.	**.428***	**.434***	**-.415***
		Sig. (2-tailed)	.	.013	.012	.016
		N	.	33	33	33
	BSQ Emotion Regulation	Correlation	.	.	**.454****	**-.725****
		Sig. (2-tailed)	.	.	.008	.000
		N	.	.	33	33
	CBQ Rothbart Effortful Control	Correlation	.	.	.	**-.376***
		Sig. (2-tailed)	.	.	.	.031
		N	.	.	.	33

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table 5

Temperament x Acoustic Correlations			Mean F0	Range of F0	Mean Jitter	Range of Jitter
Spearman's rho	BSQ Emotional Reactivity	Correlation Coefficient	.037	.222	**.468****	**.463****
		Sig. (2-tailed)	.839	.213	.006	.007
	BSQ Attention Regulation	Correlation Coefficient	-.221	.016	**-.465****	-.245
		Sig. (2-tailed)	.216	.929	.006	.170
	BSQ Emotion Regulation	Correlation Coefficient	.303	-.244	**-.551****	**-.469****
		Sig. (2-tailed)	.086	.172	.001	.006
	CBQ Rothbart Effortful Control	Correlation Coefficient	.112	.247	-.299	-.284
		Sig. (2-tailed)	.534	.166	.091	.110
	CBQ Rothbart Negative Affect	Correlation Coefficient	-.227	.286	**.566****	**.462****
		Sig. (2-tailed)	.204	.107	.001	.007

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 1

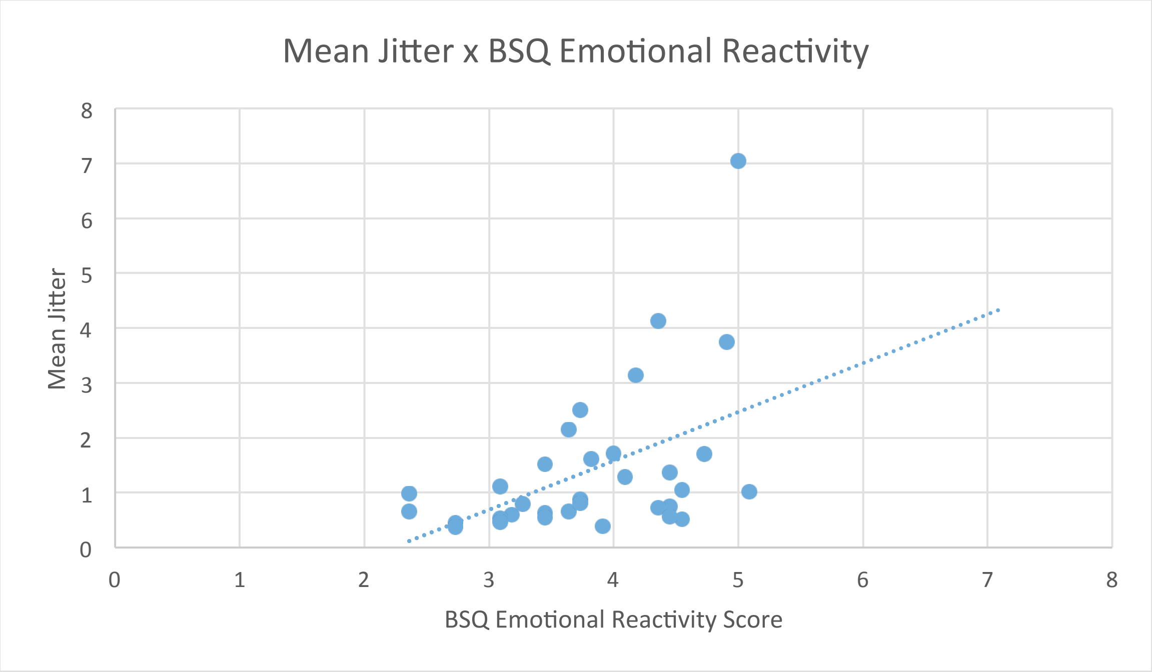


Figure 2

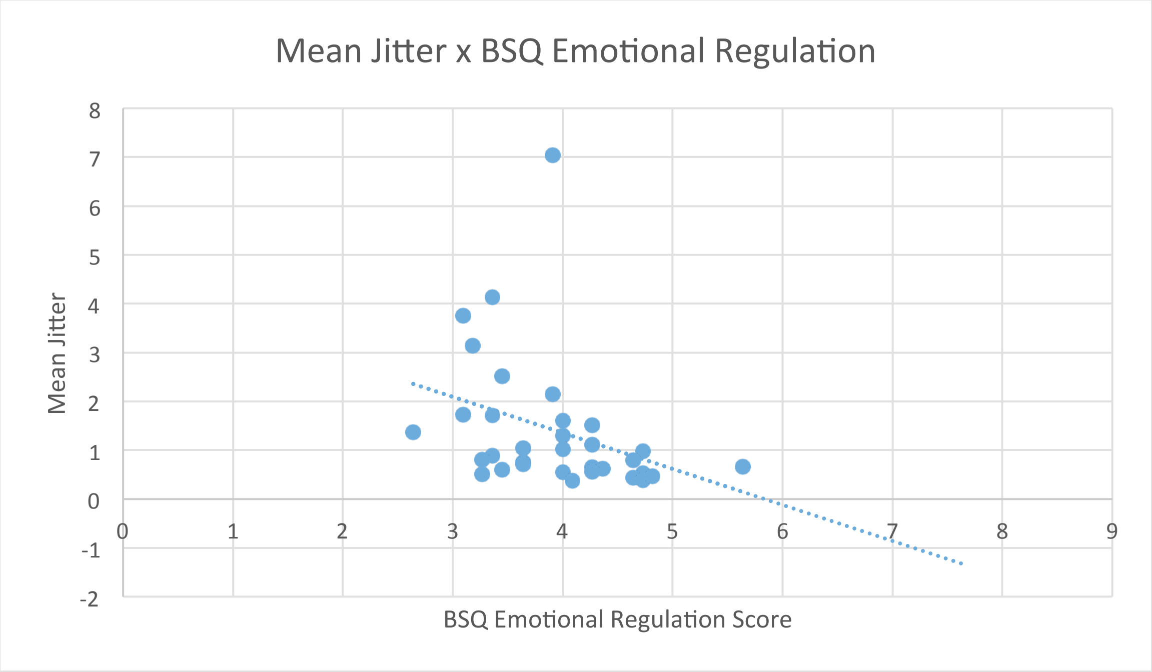


Figure 3

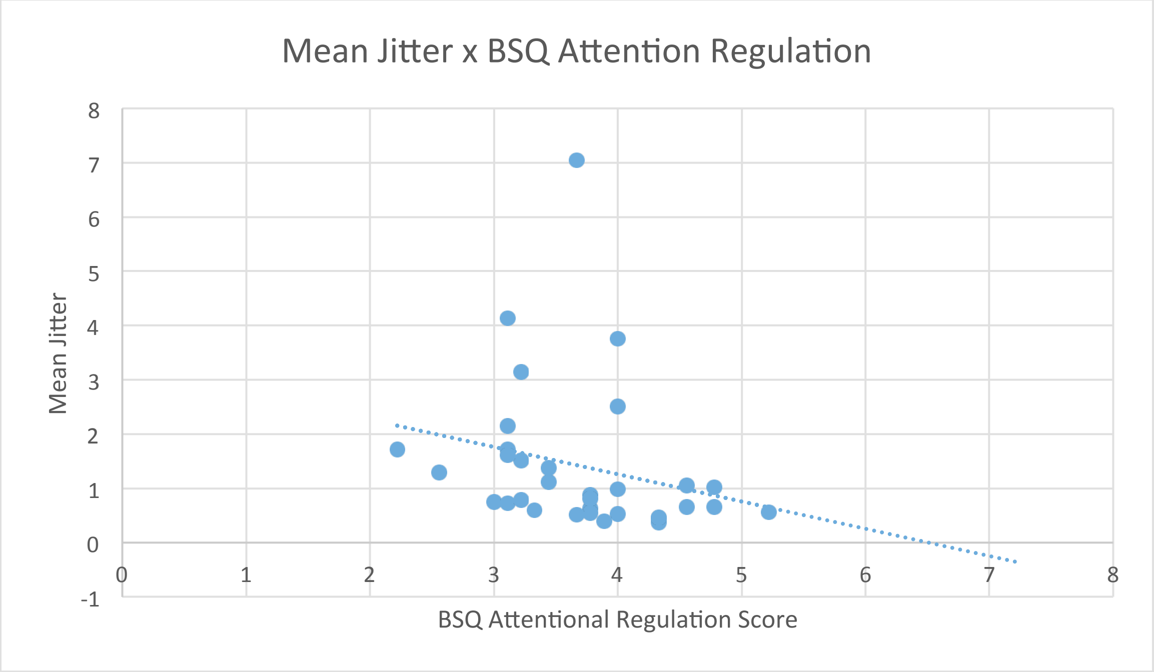


Figure 4

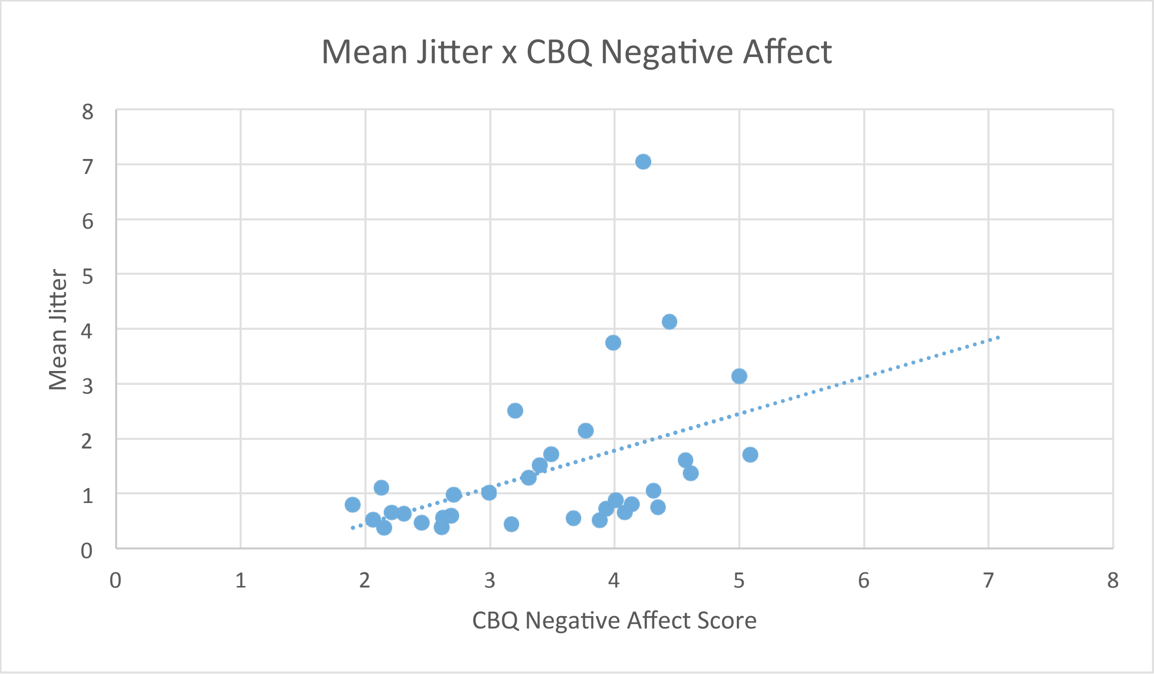


Figure 5

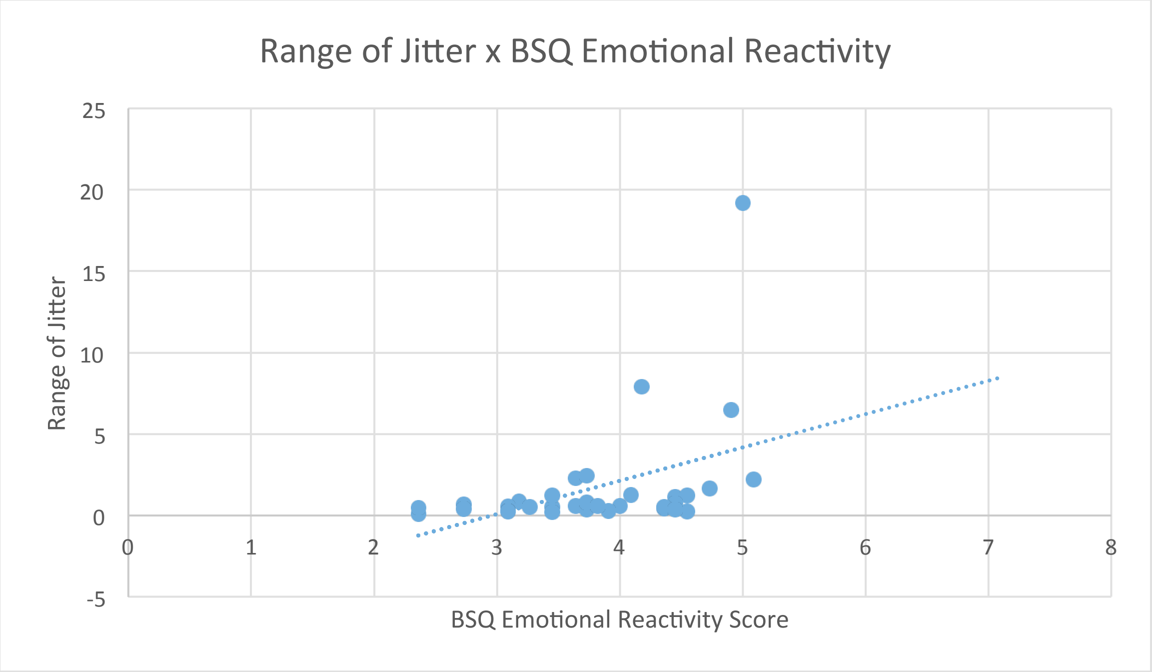


Figure 6

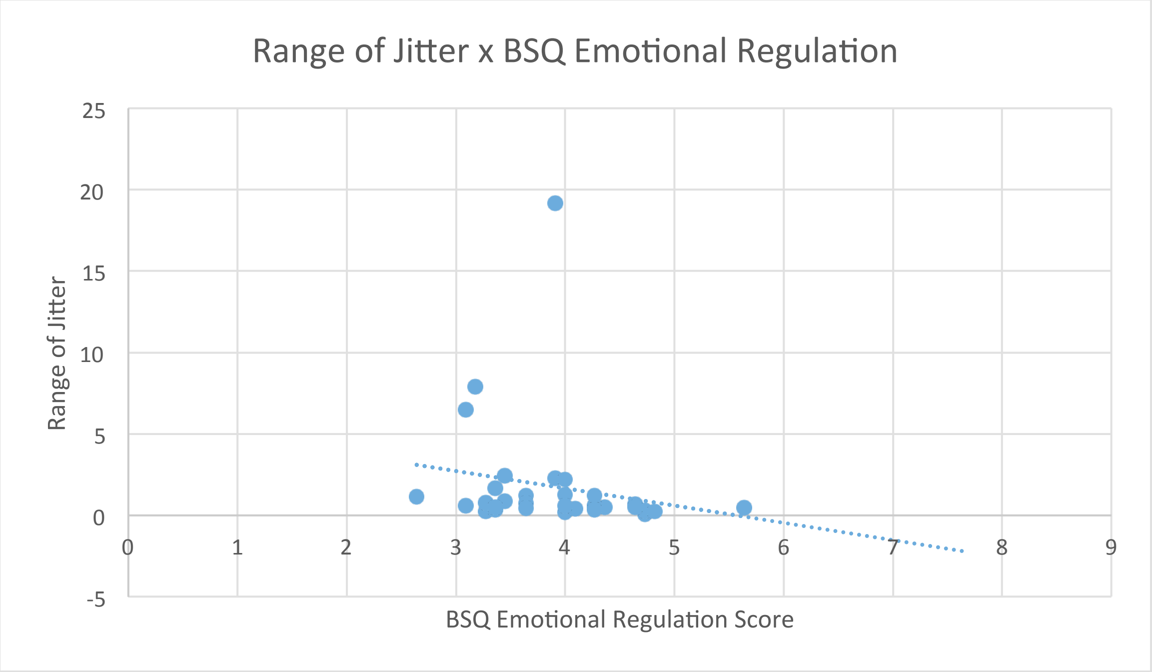
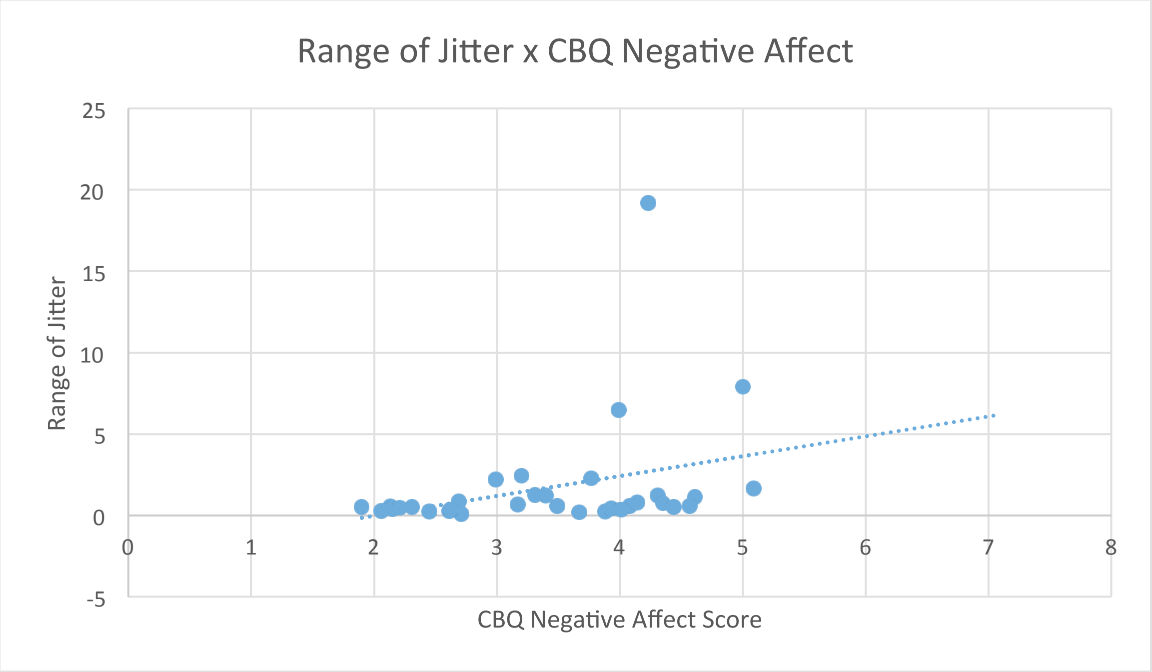


Figure 7


Range of Jitter x CBQ Negative Affect
Range of Jitter
25
20
15
10
5
0
-5
0
1
2
3
4
5
6
7
8
CBQ Negative Affect Score