Acoustic Correlates of Speech Naturalness in Post-Treatment Adults Who Stutter

Jacqueline Schmitt
Acknowledgements

Dr. Stephen Tasko
Dr. Helen Sharp
Lee Honors College
What is Stuttering?

Definition
- Interruption in forward flow of speech (Van Riper, 1982)
- Part or whole word repetitions
  - (bu-bu-but or and-and-and)
- Prolongations
  - (Mo—mmy)
- Blocks of airflow before or during a sound or word production (Yairi & Seery, 2015).

Prevalence
- 1% of general population (Yairi & Seery, 2015)
Treatment

- Fluency Shaping Treatment
- Effects of treatment (Yairi & Seery, 2015)
- Post-treatment concerns of speech naturalness
Speech Naturalness Rating Procedures

- **1-9 Interval Rating Procedures**
  - Martin et al, 1984
  - Metz et al, 1990
  - Onslow et al, 1992
  - Tasko et al, 2007

- **Direct Magnitude**
  - Metz et al, 1990

- **Agreement Ratings**
  - Schaeffer and Eichorn, 2001
  - Schaeffer, 2006
Acoustic Characteristics Influencing Naturalness

- **Vowel prolongation**
  - Vowels prolonged at more than 150% were seen by 83% unnatural (Schaeffer & Eichorn, 2001)

- **Sentence Duration**
  - Utterances prolonged at 210% (Shaeffer, 2006)
  - Correlated with picture description task \( r = 0.637 \) (Metz et al, 1990)

- **Voice Onset Time**
  - Correlated with oral reading \( r = 0.679 \) (Metz et al, 1990)
Acoustic Characteristics Influencing Naturalness

- Healey (1982)
- Studied variation of SFF two utterance types.
  - Mean SFF, SFF standard deviation, and SFF range
  - SFF standard deviation of adults who did not stutter were greater than those who did.
  - Adults who stutter had more restricted pitch ranges, less variance, more monotone.
  - Differences made them distinguishable from fluent speakers
Role of Fundamental
Frequency in Speech
Naturalness
Speaking Fundamental Frequency (SFF)

- What is SFF?
  - Rate at which vocal folds approximate together
  - Affects prosody of speech
Research Question

- Is there an association between measures of speaking fundamental frequency variation and listener ratings of speech naturalness?
Method
Participants

- Participants drawn from the Walter Reed - Western Michigan University Stuttering Database (43 adults who stutter & 43 healthy controls)
- Began with 34 adults who participated in the Walter Reed Stuttering Treatment Program
- 29 Adults (27 male, 2 female)
  - 3 excluded due to missing records
  - 1 excluded due to strong accent
  - 1 excluded due to poor audio quality of sample
Treatment Program

- One month
- Group based
- Fluency shaping
  - Three part program
Stuttering Severity Measures

O Severity ratings before and after treatment
OMonologue video
OSSI-3 judged by two speech-language pathologists
OConsensus format
Speech Naturalness Rating Procedure

- 1-9 interval rating scale
- Three groups of raters
- One minute segment from monologue before and after treatment
- Mean rating across samples was recorded
Sample Selection

- Rating based on naturalness of monologue
- Analysis of reading samples was chosen
Speaker Fundamental Frequency Analysis

Audio files were extracted from video recordings using Audacity.

Examined in TF-32 software.

SFF contours computed.

Thousands of f0 moments and the frequencies at those moments.
Speaker Fundamental Frequency Analysis

- Editing procedure completed to eliminate the discontinuities in the pitch trace
- Halving or doubling of SFF due to misinterpretation of waveform
Speaker Fundamental Frequency Analysis

- Custom written Matlab program
  - Additional edits of outliers

- SFF measures based on the several thousand f0 measurements were derived for each speech sample.
  - Mean SFF (Hz),
  - SFF standard deviation (in Hz and in semitones),
  - SFF interquartile range (in Hz and in semitones).
  - Method described in Baken (1987) was used to convert from Hz to semitones

- Descriptive statistics of behavioral and acoustic measures were generated
- Simple correlations were performed between acoustic measures and behavioral measures
Results
Speaker Fundamental Frequency Analysis

O Procedure reliability strong

<table>
<thead>
<tr>
<th>Sample-Rater Initials</th>
<th>Mean</th>
<th>Std</th>
<th>Std_St</th>
<th>IQR</th>
<th>IQR_St</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample 1 - CJ</td>
<td>106.90</td>
<td>13.73</td>
<td>2.09</td>
<td>15.96</td>
<td>2.60</td>
</tr>
<tr>
<td>Sample 1 - JS</td>
<td>107.00</td>
<td>13.87</td>
<td>2.11</td>
<td>15.98</td>
<td>2.60</td>
</tr>
<tr>
<td>Sample 2 - CJ</td>
<td>141.90</td>
<td>13.52</td>
<td>1.57</td>
<td>15.71</td>
<td>1.92</td>
</tr>
<tr>
<td>Sample 2 - JS</td>
<td>141.87</td>
<td>13.63</td>
<td>1.58</td>
<td>15.87</td>
<td>1.94</td>
</tr>
<tr>
<td>Sample 3 - CJ</td>
<td>111.94</td>
<td>15.98</td>
<td>2.31</td>
<td>17.86</td>
<td>2.76</td>
</tr>
<tr>
<td>Sample 3 - JS</td>
<td>112.19</td>
<td>15.05</td>
<td>2.179</td>
<td>17.63</td>
<td>2.72</td>
</tr>
</tbody>
</table>
# Behavioral & Acoustic Measure Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th>PRE SSI</th>
<th>Pre-tx Nat</th>
<th>PST SSI</th>
<th>Post_Tx Nat</th>
<th>Mean (Hz)</th>
<th>Std (Hz)</th>
<th>Std (St)</th>
<th>IQR (Hz)</th>
<th>IQR (St)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Average</strong></td>
<td>24.82</td>
<td>4.215</td>
<td>4.215</td>
<td>3.573</td>
<td>128</td>
<td>17.5</td>
<td>2.21</td>
<td>21.78</td>
<td>2.9</td>
</tr>
<tr>
<td><strong>Minimum</strong></td>
<td>17</td>
<td>1.2</td>
<td>0</td>
<td>1.65</td>
<td>102</td>
<td>11.16</td>
<td>1.53</td>
<td>9.75</td>
<td>1.36</td>
</tr>
<tr>
<td><strong>Maximum</strong></td>
<td>35</td>
<td>7.4</td>
<td>20</td>
<td>6.14</td>
<td>208</td>
<td>31.97</td>
<td>3.58</td>
<td>47.64</td>
<td>4.86</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td>18</td>
<td>6.2</td>
<td>20</td>
<td>4.49</td>
<td>106</td>
<td>20.81</td>
<td>2.05</td>
<td>37.89</td>
<td>3.5</td>
</tr>
</tbody>
</table>
SSI-3 & Speech Naturalness Ratings

\[ r = -0.26 \]
\[ P = 0.17 \]
## Correlation Coefficients of Acoustical Measurements, Speech Naturalness, & Stuttering Severity

<table>
<thead>
<tr>
<th></th>
<th>Pre tx SSI</th>
<th>Pre tx Nat</th>
<th>Post tx SSI</th>
<th>Post tx Nat</th>
<th>Mean (Hz)</th>
<th>Std (Hz)</th>
<th>Std_St</th>
<th>IQR (Hz)</th>
<th>IQR_St</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre tx SSI</td>
<td>--</td>
<td>.674</td>
<td>-.1</td>
<td>.26</td>
<td>.153</td>
<td>.02</td>
<td>-.07</td>
<td>.06</td>
<td>-.11</td>
</tr>
<tr>
<td>Pretx Nat</td>
<td>--</td>
<td>.02</td>
<td>.41</td>
<td>.259</td>
<td>.01</td>
<td>-.23</td>
<td>.12</td>
<td>-.1</td>
<td></td>
</tr>
<tr>
<td>Post tx SSI</td>
<td>--</td>
<td>-.26</td>
<td>.272</td>
<td>-.14</td>
<td>-.37</td>
<td>-.12</td>
<td>-.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post tx Nat</td>
<td>--</td>
<td>-.11</td>
<td>-.01</td>
<td>.09</td>
<td>.01</td>
<td>.13</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Mean SFF and Speech Naturalness

$r = -0.11$
$P = 0.56$
Standard Deviation of Semitones

\[ r = 0.09 \]

\[ P = 0.64 \]
Interquartile Range of Semitones

$r = .09$

$P = .47$
## Correlation Coefficients of Acoustical Measurements, Speech Naturalness, & Stuttering Severity

<table>
<thead>
<tr>
<th></th>
<th>Pre tx SSI</th>
<th>Pre tx Nat</th>
<th>Post tx SSI</th>
<th>Post tx Nat</th>
<th>Mean (Hz)</th>
<th>Std (Hz)</th>
<th>Std_St</th>
<th>IQR (Hz)</th>
<th>IQR_St</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre tx SSI</strong></td>
<td>--</td>
<td>.67</td>
<td>-.1</td>
<td>.26</td>
<td>.15</td>
<td>.02</td>
<td>-.07</td>
<td>.06</td>
<td>-.11</td>
</tr>
<tr>
<td><strong>Pretx Nat</strong></td>
<td>--</td>
<td>.02</td>
<td>.41</td>
<td>.25</td>
<td>.01</td>
<td>-.23</td>
<td>.12</td>
<td>-.1</td>
<td></td>
</tr>
<tr>
<td><strong>Post tx SSI</strong></td>
<td>--</td>
<td></td>
<td>-.26</td>
<td>.27</td>
<td>-.14</td>
<td>-.37</td>
<td>-.12</td>
<td>-.32</td>
<td></td>
</tr>
<tr>
<td><strong>Post tx Nat</strong></td>
<td>--</td>
<td></td>
<td></td>
<td>-.11</td>
<td>-.01</td>
<td>.09</td>
<td>.01</td>
<td>.13</td>
<td></td>
</tr>
</tbody>
</table>
Stuttering Severity & IQR Semitones

$r = -0.37$

$P = .08$
Stuttering Severity & Standard Deviation of Semitones

$r = -0.32$
$P = 0.04$
Discussion
Discussion

- Variability measurements vs naturalness ratings
- No relationships found
- Slight negative relationship between stuttering severity standard deviation and interquartile range of semitones
- The higher stuttering severity, the lower the range of SFF.
Limitations

- Reading sample vs naturalness ratings
- Variation measurements
- OSSI – 3 variability
- Amount of participants
  - Gender Variation
  - Literacy skill
- No typically speaking control group
Future Directions

- Method of measurement
  - SFF contour onset of utterances
- New data set
  - Better audio quality
  - Larger participant pool
  - Control group
Questions