Comparing Real-time and transcript-based measures of counting and categorizing stuttering

Jennifer Jackson\textsuperscript{a} & Farzan Irani\textsuperscript{b}

\textsuperscript{a}Taft ISD, \textsuperscript{b}Texas State University – San Marcos

**ABSTRACT**

The purpose of this study is to determine the reliability of a transcript based disfluency measurement compared to an online disfluency measurement system. A trained graduate student clinician analyzed five speech samples obtained from the Stuttering Measurement System website (http://sms.id.ucsb.edu) using both methods. Data was analyzed for reliability across methods.

**METHODS**

**Data**

The speech samples used for analyses in this study were obtained for the Stuttering Measurement System (SMS) homepage (http://sms.id.ucsb.edu) with permission from Dr. Roger Ingham.

**Training**

The first author completed a three-step training program comprised of 1) education/review, 2) identification of disfluencies, and 3) practicing each measuring method.

**Procedures**

The first author analyzed two samples each week using either procedure. Care was taken to ensure the same speech sample was not analyzed twice in one week.

**Results**

A number of studies have examined intra-judge and inter-judge reliability of stuttering disfluency counts using a variety of techniques including transcribing the speech sample verbatim (Rustin, Botterill, & Kelman, 1996), analyzing speech disfluencies using a real-time counting and categorizing of dysfluencies (Conture, 1990), and the use of a time-interval measurement of stuttering, which does not include a specific count of disfluency types, only a percent of total disfluencies in a speech sample (for example, Cordes, Ingham, Frank, & Ingham, 1992).

**Background**

It can be argued that categorizing the type of disfluencies might not be of the same importance (Einarsdottir & Ingham, 2005); however, when using a stuttering management approach, clinicians often teach clients to modify disfluencies to reduce tension and struggle associated with stuttering, thereby producing disfluency free of any tension and struggle (Yairi & Seery, 2011). Keeping the goal of a stuttering management approach in mind, it might be important for clinicians to categorize speech disfluency types.

**Analysis**

Descriptive Data. Data gathered from the analyses includes a count of the number of syllables and the total number of disfluencies categorized by type of disfluency. The disfluencies were further categorized as either More Typical or Less Typical types of disfluencies (Yairi & Seery, 2011).

Correlation Measures. Similarities and differences between the two measurement techniques when used by the same clinician were completed using direct comparisons of frequency counts from each technique; a Pearson Product Moment Correlation (PPMC), and Intraclass Correlation Coefficient (ICC).

**Conclusions**

Results indicate a statistically significant correlation between the transcript-based and online method of counting and categorizing disfluencies.

The descriptive data indicates that differences between the two measures exists. The correlation between specific disfluency types was lower than disfluency counts. This supports the results by Yaruss et al. (1998).

A primary weakness of this study is the small sample size (n=5). Further, the study only explored correlation between methods for the same judge. Future studies should determine inter-judge and intra-judge reliability separately for each method.

**REFERENCES**


**RESULTS**

Figure 1 provides descriptive statistics for %SS, %MT, and %LT. Overall, the difference between the two methods (online analysis vs transcript based) was found to be less than 1%.

Figure 2 provides descriptive statistics for specific types of more typical disfluencies. Data indicates a difference of less than 1% between methods for all three.

Figure 3 provides descriptive data for percentage of less typical types of disfluencies. The difference between methods for word repetition and blocks was found to be less than 1% and for prolongations it was found to be less than 2%.

Correlation analyses using a PPMC and ICC to determine the correlation indicate a strong correlation between the two methods (r=.99, p<.01; ICC=.989, p<.01).

**CONCLUSIONS**

Results indicate a statistically significant correlation between the transcript-based and online method of counting and categorizing disfluencies.

The descriptive data indicates that differences between the two measures exists. The correlation between specific disfluency types was lower than disfluency counts. This supports the results by Yaruss et al. (1998).

A primary weakness of this study is the small sample size (n=5). Further, the study only explored correlation between methods for the same judge. Future studies should determine inter-judge and intra-judge reliability separately for each method.

**CONTACT INFORMATION**

Farzan Irani, PhD., CCC-SLP; irani@txstate.edu
Department of Communication Disorders
Texas State University – San Marcos