A STUDY OF THE INTERRELATIONSHIPS BETWEEN THE TIME-INTERVAL DISFLUENCY TEST AND SANDER'S DISFLUENT WORD INDEX

by

CAROL COBURN HAMMOND

B. A., Kansas State University, 1972

A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

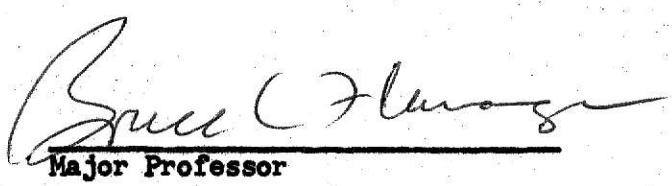
MASTER OF ARTS

Department of Speech

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1973

Approved by:



Major Professor

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

The author wishes to express her gratitude to Dr. Bruce C. Flanagan for his guidance in preparation of this thesis. Special acknowledgement is given to Dr. James B. Lingwall of the University of Kansas for his cooperation in making subjects and facilities available for obtaining data. The author wishes to thank the students in Stuttering, 283-834, who devoted time and effort in analyzing the data contained in this research.

A special thanks is extended to my parents, Mr. and Mrs. C. M. Coburn Jr., for their continuous encouragement and financial support throughout my college experience.

LIST OF TABLES

CHAPTER I

INTRODUCTION

Rating scales have long been the predominant means of determining the severity of stuttering and assessing progress in therapy. The first rating scale developed was by Lewis and Sherman (1951). The Lewis and Sherman Scale of Stuttering Severity consists of recorded samples of stuttered speech, each sample representing a point on an equal appearing interval scale. Berry and Silverman (1971) did a study on the equality of intervals on the Lewis-Sherman Scale and reported the intervals on equal appearing scales are "equal appearing" because the observers are instructed to make their ratings in this manner. They suggest it is not safe to assume that the intervals on equal appearing scales are subjectively equal.

Starbuck (1954) developed a filmed audiovisual scale in which samples of stutterers in the act of stuttering were arranged sequentially to represent differing degrees of severity. Starbuck showed that five factors varied significantly with each successive step on his scale. The five factors were total words spoken in a given time, number of blockings, facial grimances, accompanying limb movement, and eye shift or eye blink. Van Riper (1971), in comparing the Lewis and Sherman Scale with the Starbuck Scale, stated that for actual clinical use both tests proved difficult to use primarily because it is difficult to match one stutterer against another.

Probably the most widely used of all rating scales is the Scale for Rating the Severity of Stuttering by Johnson, Darley, and

Spriesterbach (1963). Seven scale values are given with descriptive legends for each. Although the authors strongly emphasize the limitations of this scale and term it a "rough measure only", many speech clinicians use it for diagnosis and measuring therapeutic progress.

In an attempt to examine various rating scales Cullinan, Prather and Williams (1963) used seven different rating scales to evaluate the severity of tape recorded samples of stuttering. They found that with only one judge, no one method was reliable enough to serve as an adequate predictor of severity for an individual stutterer, and none of the seven methods seemed to be any better than the others. Young (1970) reported that scale values of equal appearing rating scales are affected by a variety of extraneous conditions. In particular, anchoring and sequence effects are known to result in predicting observer bias.

More complex rating scales have been developed recently by Riley (1972) and Wertheim (1972). Riley's measurement is called the Stuttering Severity Instrument (SSI). The SSI is designed for clinical and research use and yields a single numerical representation of severity with a range of 0-45. Frequency, duration, and physical concomitant are scored on a point system by the evaluator. Wertheim is presently researching a stuttering measurement that would provide information on the situational stability of the disorder, its qualitative pattern, its severity, and the interrelationship between these parameters for a given individual. Both Riley and Wertheim call their new measurements objective; however, both require the summation of three observer ratings on the part of one or more listeners.

Allyon and Azrin (1968) have suggested that distortion and error can result from interpretive characterization of performance, utilizing rating scales. They recommend describing the behavior in specific terms that require a minimum of interpretation. In this way, the observers would avoid reading their interpretations into the data by forcing them to attend to and record only in terms of the physical aspect of the response. Realizing a need for more descriptive procedures that would bypass the relative subjectivity of rating scales, several methods of measuring disfluencies have been established.

Johnson (1961) developed a method of measuring disfluencies known as the Iowa Speech Disfluency Test. The speech behavior classified as disfluencies were interjections of sounds, syllables, words or phrases, part word repetitions, phrase repetitions, revisions, incomplete phrases, broken words, and prolonged sounds. In this study measurements were made of rate and disfluency in samples of oral reading and speaking. Subjects consisted of 100 male and 100 female adult speakers, of whom 50 in each group were classified as stutterers and 50 as nonstutterers. Each subject performed three speaking tasks. These tasks were the Job Task, responses to Thematic Apperception Test Cards, and Oral Reading. Disfluencies were identified in each case from a verbatim transcript while listening to a play-back of the recording. The recording was replayed as often as necessary to insure an accurate identification of disfluencies. The subjects' verbal output in each task was defined as the number of words spoken. Speaking and reading rates were computed in terms of words per minute by calculating the ratio of verbal output to reading or speaking time as measured by a stop watch. A computation was made of the number of instances of

each type of disfluency per 100 words for each of the three speaking tasks, giving a Disfluency Category Index. The formula used to compute this index was (ND/NW) 100=Disfluency Category Index.

> ND represents total number of instances of disfluencies of the designated type in the speech sample of the subject.
>
> NW represents the number of words (verbal output) of the subject for the sample.

A disfluency index for each task was determined for each subject by obtaining the sum of the subject's category indexes for the task.

Inter and intra-examiner reliability was established through three separate investigations. Duffy (1957) obtained data from two observers who listened to 12 recorded speech samples of female stutterers and recorded disfluencies using the eight categories. He correlated the results of the two observers in identifying disfluencies in each category and obtained co-efficients of correlation ranging from .90 to .99. Young (1961), in order to establish intra-examiner reliability, analyzed 10 tape recorded speech samples of adult male stutterers twice, with intervening periods of from three weeks to two months with reference to five types of disfluent categories. Indexes ranged from .91 to 1.00, for the 10 samples, with an index of agreement of .97 for all 10 samples combined. Sander (1961) established intra-examiner reliability by making two analysis of the recording of 12 Job Tasks and 12 reading tasks performed by male stutterers. The analysis were separated by at least a one month interval. Percentage of agreement for total disfluency count between the first and second scoring was .96 for both Job Task and reading tasks combined. Johnson presents his data including the following: verbal output in the Job and TAT tasks, measures of time spent on each task, ranges and deciles of

distributions of oral reading and speaking rates for the three tasks, measures of disfluency, and co-efficients of correlation among the three tasks for measure of each of the disfluency variables and of rate, in terms of tables. Johnson utilizes these findings as normative and comparative data for the Iowa Speech Disfluency Test, with respect to rate and disfluency in the speech and oral reading of adult male and female stutterers and nonstutterers.

Young (1961) has also developed a method of measuring disfluencies. Disfluencies were classified as interjections, part-word repetitions, word and phrase repetitions, prolongations, and revisions. Subjects consisted of 50 college age male stutterers. Recordings were made of the Job Task to secure samples of speech.

First a verbatim transcription was made and then the disfluencies were identified and classified. The tape was replayed as often as necessary to obtain sufficient accuracy. To estimate intra-examiner reliability in making the fluency analysis, 10 tapes were selected at random and a second fluency analysis of each was performed with a minimum of two weeks and a maximum of three months separating the analyses. Intra-examiner reliability in identifying total disfluencies per sample ranged from .91 to 1.0, while for all 10 samples combined, the index was .97. Young presents his data, including a summary of the five measures of disfluency, a frequency count for each type of disfluency, a measure of speaking time, and total frequency of all types of disfluency, in terms of mean frequency, range, and standard deviation.

Sander (1961) conducted a study concerned with establishing the test-retest reliability for both reading and speaking tasks of the Iowa Speech Disfluency Test administered to a group of stutterers over

an interval of 24 hours. Disfluencies were defined as ones that involved a prolonged sound or a broken utterance, a word interrupted by an interjection, or a repetition of sound, syllable or word. Sander's Disfluent Word Index involves the eliciting of a recorded sample of reading and spontaneous speech. Subjects consisted of 40 stutterers, all but one participating in therapy at the University of Iowa Speech Clinic. Thirty-four of the stutterers were male, six were female. They ranged in age from 17 to 37 years with a mean age of 22.6 years. There was a 24 hour interval between the initial and subsequent administration of the reading and speaking tasks. The speaking situation was held constant for both days. The recording was then replayed as often as necessary to determine and classify the speaker's disfluencies. Rate of utterance for the 300 word passage and the 250 word speaking task was also determined.

Sander established intra-examiner reliability by rescoring 12 reading and speaking tasks one month after the original analysis. He found co-efficient of agreement to be .96 for total disfluent words. This result was for total number of disfluencies only and did not indicate the extent of agreement for the individual disfluency categories. Sander presented the means, standard deviations, and ranges for total disfluencies, disfluent words, and time in a table. He ran a test-retest Pearson product-moment correlation to determine reliability of the subject's behavior. Results indicated that over the 24 hour period, half of the subjects showed no shifts in decile ranking, and over four-fifths of the subjects showed changes not greater than one decile. These results indicate rather high temporal reliability, suggesting the feasibility of using such a process to measure disfluencies.

Each of the aforementioned studies, Johnson, Young, and Sander, utilize a method consisting of measuring rate of utterance and classifying disfluencies into four to eight classifications. That these methods are time consuming and tedious is supported by the clinical experience of the current author, which tends to indicate that these measures take from two to four hours to complete and analyze. Johnson himself concedes, "Much time is needed, often several hours per subject, to obtain accurate transcriptions and identifications of the disfluencies in individual speech samples." The length of time required to administer and score these indexes (analysis and categorization of disfluencies for 300 word speaking samples and computation of rate of utterance), would suggest a need for a more efficient procedure that affords equivalent reliability.

Another procedure for obtaining a quantitative measure of the frequency of disfluencies is tabulating the number of times that a specified behavior occurs in a given time period. However, when a continuous measure of the frequency of specific behaviors is desired, it is useful to record the presence or absence of these behaviors in discrete time intervals.

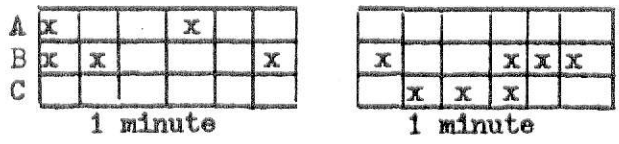
Bijou, Peterson, and Ault (1968) have suggested a method of recording behavioral events in field situations consisting of registering the frequencies of occurrences and non-occurrences within a time interval. The observer makes a mark and only one mark, in each time interval in which the response occured. In this procedure, it is obvious that the maximum frequency of a response is determined by the size of the time unit selected, i.e. if a 10 second interval was utilized, the maximum rate would be six responses per minute. If the

frequency of occurrence of the desired behavior is high, smaller time intervals should be employed to obtain a high correspondence between the recorded and actual frequencies of occurrences. This procedure has been effectively used in recording spontaneous speech, tantrum behavior, and attending-to-work behavior.

Johnston, Kelly, Wolf, and Harris (1968) have also developed a time series analysis procedure where the observation period is broken up into 10 second intervals. For each interval, the absence or presence of the behavior being studied is recorded. Since this procedure simply records whether or not a specific behavior occurred in an interval and ignores the number of times it occured in the interval, it is important that the intervals are short enough so that the specific behavior should be used to determine the size of the interval. However, with manual recording, it is impractical to use intervals shorter than 10 seconds.

To better illustrate the Time Series Analysis Method, a visual representation of a series of 10 second intervals and the scoring involved is presented.

Behavior A: Laughing
Behavior B: Talking
Behavior C: Crying

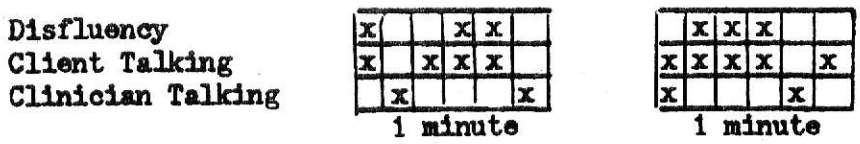


The behavior of laughing occured during two of the 10 second intervals out of 12 ten second intervals.

The behavior of talking occurred during seven 10 second intervals out of 12.

The behavior of crying occurred during three 10 second intervals out of a total of 12 ten second intervals.

Utilizing this procedure, a method of measuring disfluencies has recently been developed known as the Time Interval Disfluency Test. This test involves the eliciting of a recorded sample of speech in three speaking situations: oral reading, conversation, and monologue. At 10 second intervals a beep is recorded simultaneously with the speech sample. For each 10 second interval, the listener indicates on a score sheet if the client is speaking, if he is disfluent or if the interviewer is speaking. The disfluency index is expressed in terms of the percent of time the subject was speaking and the percent of time he was disfluent. The following is a visual representation of the Time Interval Disfluency Test.



$$\frac{\text{Client Talking}}{\text{Total Time Available}} = \% \text{ intervals of time client was talking}$$

$$\frac{\text{Disfluencies}}{\text{Client Talking}} = \% \text{ disfluent intervals}$$

$$\frac{\text{Clinician Talking}}{\text{Total Time Available}} = \% \text{ intervals of time clinician was talking}$$

The client was talking 75% of the time.
The clinician was talking 33.3% of the time.
The client was disfluent 66.6% of the time.

This procedure of measuring disfluencies eliminates the tedious counting of each disfluent word in a speaking sample of 250-300 words, and the computation of rate of utterance. There is no need for a verbatim transcription to be written or time consuming categorization of

disfluency types. The listener records whether or not the speaker was speaking or disfluent during that particular 10 second interval.

Realizing the need for a valid and effecient method of disfluency analysis, this research attempts to study the relationships between the TIDT and an established method of disfluency analysis, the Sander's Disfluent Word Index. The purpose of this study is to obtain the coefficient of correlation between the TIDT and Sander's Disfluent Word Index. Three speaking tasks are involved in the TIDT and a measure of rate of speaking time for each of the three tasks. The first task is oral reading, where the subject is asked to read from cards containing paragraphs on various topics. The second task, monologue, involves the subject talking about various colored pictures presented to him. The third task, conversation, involves dialogue between subject and examiner. Percentage of disfluent words and rate for each task will be figured as shown in the examples below.

$$\frac{\text{Disfluencies}}{\text{Client Talking}} = \% \text{ disfluent intervals}$$

$$\frac{\text{Client Talking}}{\text{Total Time Available}} = \% \text{ intervals of time client was talking}$$

Sander's first speaking task is oral reading, where the subject reads a 300 word passage. The second speaking task is the monologue (Job Task), where the subject must tell about his future occupation for at least a 250 word sample. Sander's percentage of disfluent words is obtained by counting the number of disfluent words identified and dividing by the total number of words.

Oral Reading

$$\frac{\text{Disfluent words}}{\text{300 total words}} = \% \text{ disfluent words}$$

Job Task

$$\frac{\text{Disfluent words}}{\text{250 total words}} = \% \text{ disfluent words}$$

Sander computes the rate of speaking time in terms of counting total number of seconds for each speaking task.

These disfluency scores and rate of speaking time for each task of the TIDT and Sander's Disfluent Word Index will be computed for each subject. A Pearson Product-Moment correlation will be used to obtain these coefficients of correlations. These correlations will be obtained between the five speaking tasks and the corresponding five measurements of rate. These correlations should reveal if the TIDT is a valid test as compared to Sander's Disfluent Word Index. Other information, such as time and preparations necessary to administer both procedures and examiner problems in scoring, will be discussed and compared.

CHAPTER II

METHOD

Selection of Subjects

Nine male stutterers served as subjects. Their ages ranged from 19-42 years. The mean age of the group was 24.8 years. Seven of the subjects were obtained from the Kansas State University Speech Clinic and two were obtained from the University of Kansas Speech Clinic. All of the subjects were receiving speech therapy for stuttering.

Recording Procedures

An Ampex two-channel tape recorder, model number AG 500, was used to transmit the 10 second beeps (1K Hz tones) simultaneously on to the speech sample as it was being recorded. All recording equipment was in an adjourning room to the experimental room and was out of sight of the subject. An over-head microphone was the only visible sign of recording equipment.

Each subject was seated at a table in a sound treated room with an over-head microphone positioned approximately 18 inches from his mouth. The experimenter, seated next to the subject, asked identifying information such as name, age, address and phone. This was done to accustom the subject to the experimental situation. The recording level for each speaker was adjusted during these introductory remarks.

In the TIDT recording procedures for the oral reading situation, the subject was handed 5" x 7" index cards and asked to read them aloud

as he normally would. Each card contained a paragraph exerpt from an encyclopedia. Each card contained information on a different topic. These cards were presented in a random order for each subject. The subject was instructed to read until he was asked to stop. (See Appendix A). A stop watch was used to time the four minute sample.

In the monologue speaking task the subject was handed 5" x 7" index cards with a single colored photograph on each one. These pictures were taken from various magazines and chosen to elicit spontaneous speech. The pictures were presented in a different order for each subject. The subject was instructed to look at each picture, one at a time, and to talk about it. He was encouraged to discuss any detail of the picture. He was instructed to go at his own rate and say as much or as little as he wanted to on each picture. The examiner did not prompt the subject after the monologue was initiated. The subject was stopped after a four minute sample had been obtained. (See Appendix B).

The third speaking situation of the TIDT was conversation or dialogue between the speaker and examiner. Questions were asked by the examiner to elicit conversational speech. Questions that could be answered by yes or no were avoided. These questions asked information such as what the subject likes to do in his leisure time, what hobbies or interests does he have, what course has he really liked or disliked, and what places has he traveled to or would like to go. (See Appendix C). The experimenter would not only ask questions but would respond appropriately to the subject's conversation as if they were having a typical conversation. A stop watch was used to insure a four minute sample.

A speech sample was also collected the way Sander collected his speech samples so they could be analyzed according to his disfluency index and rate of utterance analysis. The first task was reading. The reading passage used by Sander was the "Test Passage for Measurement of Reading Rate" (Fairbanks 1940). This passage contained 300 words. The subject was instructed to read it aloud as he normally would. (See Appendix A).

Sander's second task, titled Job Task, was a monologue. The subject was instructed to talk for three minutes about his preferred, possible, or future job or vocation. The subject was told to describe the vocation, tell why he chose it, or anything else about it he wished. The subject was given time to organize this thought and then say when he was ready to begin. If the subject could not talk on his future vocation for three minutes he was asked to tell about past jobs. (See Appendix B). If the subject had a problem in speaking the full three minutes, leading questions were asked to encourage him to continue. The subject was asked to talk until a three minute sample or approximately 250 words had been recorded for the speaking task. In each of the five speaking tasks, for both TIDT and Sander, the order of presentation of the speaking tasks was randomly varied with each subject.

Disfluency Analysis

To accommodate both the TIDT and Sander's Disfluent Word Index the definition of a disfluency is Sander's own. A word was considered to be disfluent if it involved prolonged sounds, was classified as a broken word, was involved in sound, syllable or word repetition, or was interrupted by an interjection. Words preceeded by interjections or

involved in a phrase repetition were not counted as disfluent words. The repetition of a given word was counted as a single disfluency regardless of the number of times the word or portion of the word was repeated. Similarly, the interjection of a single word or of an entire phrase was counted as a single disfluency. This definition was held constant whether using the TIDT or Sander's disfluency analysis.

Eleven judges were used to evaluate the speech samples using both procedures, TIDT and Sander's. The judges were all graduate students in Speech Pathology and members of a class in Stuttering 283-834, which was taught Spring Semester, 1973. To familiarize them with Sander's Disfluent Word Index they received a handout and a training session. The handout consisted of the definition of a disfluent word and instructions on Sander's procedure. (See Appendix D). The judges' instructions were to count the total number of disfluent words for each task, oral reading and monologue, for each subject. The judges were instructed to listen to each speaking task three times and record the number of disfluent words they identified for each of the two tasks.

The judges were each given a transcript of each speaking task for each of the nine subjects. The transcript was given out because Sander had done this in his research and an attempt was made to replicate his procedures as closely as possible. Two duplicate tapes of each subject were made available so the judges could randomly analyze the tapes independent of one another. The training session consisted of going over the handout sheet, giving examples of the types of disfluencies to listen for, and answering any questions about the instructions or procedure.

The analysis for the TIDT procedure was done in three separate sessions involving all eleven judges simultaneously. The training procedure consisted of verbal instructions and trial runs. To insure all the judges understood the procedure, a two minute segment of the speech sample was played. The judges had to listen and score the speech sample. The judges had to show agreement of 90% with a previously scored "key" for the sample. The "key" was developed by three listeners. A mark recorded by at least two out of the three listeners was considered correct and scored on the "key" sheet. If the judges agreement was less than 90% then another two minute sample was played and scored. This was done until all eleven judges had scored at the 90% agreement level.

The three scoring sessions for the TIDT were done one week apart. Only one speaking task was scored per session. The training procedure was repeated before each scoring session. The order of scoring subjects was varied each session. The judges were instructed each session to mark only if they heard a disfluency, not how many, within a 10 second interval. They also recorded within a 10 second interval if the client talked and whether the clinician talked during this period of time. The definition of a disfluency was the same as for the Sander's analysis.

Data Analysis

The TIDT yields three direct measures of disfluency in reading, monologue, and conversational tasks. The Sander's Index yields two direct measures of disfluency, one for the reading task, the other for the Job Task, which may be considered a combination of the monologue

and conversational tasks of the TIDT. To assess the validity of these measures of the TIDT, product-moment correlations were obtained between these measures for two sets of data. The first set of data is the mean of rate of disfluency for each subject as counted by the eleven judges. The second set of data is the count of the disfluency judgements made by the experimenter. Thus estimates can be made of how well one set of disfluency data, the Sander's Index sub-tests, may be predicted from the sub-tests of the TIDT.

Besides the disfluency analysis, rate of speaking time was obtained. For the TIDT's three speaking tasks, three direct measures of speaking rate were obtained. The Sander Index yields two direct measures of rate, one for each speaking task. Product-moment correlations were obtained between these measures for two sets of data. The first set is the mean rate of speaking for each subject as recorded by the eleven judges. The second set of data is the rate of speaking for each task as determined by the experimenter. As in the disfluency analysis, estimates can be made of how well one set of data on rate of speaking time, the Sander's Index sub-tests, may be predicted from the sub-tests of the TIDT.

Scoring Reliability

The oral reading and monologue tasks of Sander's procedure was rescored for total number of disfluencies to determine the reliability of the experimenter's analysis. In each case a one month period elapsed between the original analysis and the rescoring. The formula used to establish reliability for total disfluencies was Agreement Index = $a/(a+d)$ in which a = agreements and d = disagreements (the discrepancy

between the original and rescored tasks). The coefficient of agreement obtained from this rescoring was .94. This score is for both oral reading and Job Task combined. For inter-judge reliability scores see Buetzer (1973).

CHAPTER III

RESULTS

The results of this research are reported in three parts. These parts are: 1) a comparison of Sander's (1961) sample of stutterers to the sample of stutterers observed in this study; 2) an exposition of the individual TIDT scores for the subjects; and 3) the coefficients of correlation of the percent disfluent words and time speaking (rate) between Sander's Index and the TIDT scores.

Comparison of Samples

Table 1 shows the means, standard deviations, ranges of disfluent words and time in seconds of the reading task (Task A) and the Job Task (Task B) of Sander's Disfluent Word Index for Sander's (1961) study and the present research. Scores of disfluency for both Sander and the Mean of the Judges are reported as percent of words spoken.

For Sander the mean percent of disfluent words for the reading task was 9.7 percent. For the Mean of the Judges in the present study it was 6.7 percent. The standard deviation for Sander was 12.9 percent and for the Mean of the Judges was 7.4 percent. Sander's range of disfluency was 0-61.3 percent and for the Mean of the Judges the range was .7-23.7 percent.

For the Job Task, Sander reported a mean of 15.3 percent of words spoken to be disfluent. The mean disfluent percentage of the present study was 11.7 percent. The standard deviation of Sander's research for the Job Task was 13.4 percent and for the Mean of the Judges it was 9.0 percent. The range for Sander was .4-53.6 percent and the range for the present study was 1-31.6 percent.

TABLE 1. Means, standard deviations, ranges of disfluent words, and rate of the Reading and Job Task of Sander's Disfluent Word Index for Sander (1961) and Mean of Judges (1973).

	READING (Task A)			JOB TASK (Task B)		
	Mean	SD	Range	Mean	SD	Range
% Disfluent Words						
Sander (1961)	9.7	12.9	0-61.3	15.3	13.4	.4-53.6
Mean of Judges (1973)	6.7	7.4	.7-23.7	11.7	9.0	1-31.6
Time in Seconds						
Sander (1961)	218.4	206.9	89-1028	242.4	207.3	86-1235
Mean of Judges (1973)	170.9	65.8	93-283.5	190.3	93.9	97-393.3

The time speaking is listed in total number of seconds required to complete each task. In the reading task Sander had a mean reading time of 218.4 seconds while the Mean of the Judges mean reading time was 170.9 seconds. The standard deviation for Sander was 206.9 seconds and for the Mean of the Judges was 65.8 seconds. The range in reading time for Sander extended from 89-1028 seconds. For the Mean of the Judges the range in reading time extended from 93-283.5 seconds. For the Job Task the mean speaking time for Sander was 242.4 seconds and for the Mean of the Judges was 190.3 seconds. The standard deviation of speaking time for Sander was 207.3 seconds and for the Mean of the Judges was 93.9 seconds.

Table 2 shows the means, standard deviations, range of disfluent words and rate of reading and Job Task of Sander's Disfluent Word Index for Sander (1961) and the Experimenter (1973). The scores for disfluent words are again recorded as a percentage of total words spoken.

For the reading task the percent of disfluent words for Sander's sample showed a mean of 9.7 percent and for the Experimenter was 5.7 percent. The standard deviation for Sander was 12.9 percent and for the Experimenter it was 7.3 percent. Sander's range extended from 0-61.3 percent while the range of the Experimenter extended from .3-24.0 seconds. For the Job Task the percent of disfluent words for Sander showed a mean of 15.3 percent while for the Experimenter the mean was 7.4 percent. The standard deviation for Sander was 13.4 percent while for the Experimenter it was 4.6 percent. The range of disfluent words was .4-53.6 percent for Sander and .4-14.4 percent for the Experimenter.

The time speaking for the reading and Job Task was recorded in total number of seconds. For Sander's sample the reading task showed

TABLE 2. Means, standard deviations, range of disfluent words, and rate of Reading and Job Task of Sander's Disfluent Word Index for Sander (1961) and Experimenter (1973).

	READING (Task A)			JOB TASK (Task B)		
	Mean	SD	Range	Mean	SD	Range
% Disfluent Words						
Sander (1961)	9.7	12.9	0-61.3	15.3	13.4	.4-53.6
Experimenter (1973)	5.7	7.3	.3-24.0	7.4	4.6	.4-14.4
Time in Seconds						
Sander (1961)	218.4	206.9	89-1028	242.4	207.3	86-1235
Experimenter (1973)	170.9	65.8	93-283.5	190.3	93.9	97-393.3

a mean reading time of 218.4 seconds and the Experimenter showed a mean time of 170.9 seconds. The standard deviation for Sander was 206.9 and for the Experimenter was 65.8 seconds. The range of speaking time extended from 89-1028 seconds for Sander's sample and from 93-283.5 seconds for the Experimenter. For the Job Task the mean speaking time by Sander was 242.4 seconds and 190.3 seconds was reported by the Experimenter. Sander showed a standard deviation of 207.3 seconds while the Experimenter showed a standard deviation of 93.9 seconds. Sander's range was 86-1235 seconds and the Experimenter's range was 97-393.3 seconds.

In Tables 1 and 2 the Experimenter and the Mean of the Judges showed similar results in both percent disfluent words and speaking time. Both the scores of the Mean of the Judges and the Experimenter in both tasks, disfluent words and speaking time, showed the present sample of stutterers to be milder as a group than the group of stutterers studied by Sander.

TIDT Scores for Mean of Judges and Experimenter

Table 3 shows the mean scores of eleven judges for percent of disfluent words and percent speaking time (rate) of the three TIDT speaking tasks for each of the nine subjects. The total mean for all subjects combined is listed for reading, monologue and conversation. Each of the three tasks are listed and divided into two sets of results. Under each speaking task are listed the mean score for each subject in terms of percent disfluent and percent speaking.

In the reading task (Task C) for the nine subjects there is a range of 10.2-93.5 percent disfluent words. Because the rate of

TABLE 3. Mean of eleven judges for percent of disfluent words and percent speaking (rate) of TIDT speaking tasks for each of nine subjects and total mean for all subjects for oral reading, monologue and conversation.

	READING (TASK C)		MONOLOGUE (TASK D)		CONVERSATION (TASK E)	
Subjects	% disf.	rate*	% disf.	rate*	% disf.	rate*
1	64.4	100.0	37.9	94.3	69.6	95.8
2	93.5	100.0	64.0	100.0	56.1	95.1
3	90.5	100.0	50.0	100.0	39.9	95.4
4	27.3	100.0	24.0	100.0	53.0	100.0
5	31.1	100.0	65.9	100.0	65.2	96.6
6	21.6	100.0	30.7	100.0	32.6	100.0
7	41.7	100.0	66.3	100.0	44.2	94.7
8	10.2	100.0	20.8	100.0	8.4	99.6
9	36.0	100.0	56.8	100.0	61.6	99.6
Means of all Subjects	46.3	100.0	46.3	99.4	47.8	97.7

* Rate = % intervals subject was speaking

speaking measured a reading task all subjects spoke 100 percent of the time. The total mean disfluency rate for all subjects in reading was 46.3 percent disfluent. In the monologue task (Task D) the range of scores of the subjects was 20.8-66.3 percent for disfluent words. Only one subject was below the 100 percent level for speaking time. The mean percent disfluent words for all subjects was 46.3 percent. The mean percent for speaking time was 99.4 percent. In the conversation task (Task E) the scores for the nine subjects ranged from 8.4-69.6 percent. The range of speaking time ranged from 94.7-100 percent. The total mean for all subjects was 47.8 percent disfluent words and the mean speaking time was 97.7 percent.

Comparing the three tasks to each other the total means of disfluent words for reading, monologue and conversation were 46.3, 46.3, and 47.8 percent respectively. This is remarkable because of the wide diversity of scores for each subject within the speaking task. For each subject there was a wide range of scores for each of the three tasks required of the subject. Not one subject's set of three scores fell within a 10 percent percentage points of each other. One subject's scores, subject 6, ranged within 11 percent. The greatest diversity of scores exhibited by one subject was subject 3, whose scores ranged from 39.9 percent disfluent in conversation to 93.5 percent disfluent in reading.

Table 4 shows the Experimenter's analysis for percent disfluent words and percent speaking time of the TIDT speaking tasks for each of the nine subjects. The total mean for all subjects is listed in terms of percent disfluent and percent speaking time. In the reading

TABLE 4. Experimenter's analysis for percent of disfluent words and percent speaking (rate) of TIDT speaking tasks for each of nine subjects and total mean for all subjects for oral reading, monologue and conversation.

	READING (TASK C)		MONOLOGUE (TASK D)		CONVERSATION (TASK E)	
Subjects	% disf.	rate*	% disf.	rate*	% disf.	rate*
1	70.8	100.0	30.4	91.7	56.5	95.8
2	95.8	100.0	79.2	100.0	63.2	91.7
3	83.3	100.0	54.2	100.0	43.5	95.8
4	29.2	100.0	20.8	100.0	54.2	100.0
5	29.2	100.0	75.0	100.0	69.2	95.8
6	25.0	100.0	29.2	100.0	16.7	100.0
7	41.7	100.0	62.5	100.0	21.3	95.8
8	4.2	100.0	20.8	100.0	4.2	100.0
9	41.7	100.0	54.2	100.0	70.8	100.0
Means of all Subjects	44.5	100.0	47.3	99.1	44.4	97.2

* Rate = % intervals subject was speaking

task the subjects scores ranged from 4.2-95.8 percent. The percent speaking of all subjects was 100 percent. The mean for all subjects combined 44.5 percent disfluent words. In the monologue task the scores ranged from 20.8-79.2 percent disfluent. One subject was below 100 percent in speaking time. The mean score for all subjects was 47.3 percent disfluent words. The mean percent of speaking time for all subjects was 99.1 percent. In the conversation task the disfluency scores ranged from 4.2-70.8 percent. The range of scores for percent speaking time extended from 91.7-100 percent. The mean for all subjects in terms of disfluent words was 44.4 percent. The mean for speaking time was 97.2 percent for all subjects combined.

Looking at individual subjects for the three tasks combined, only one subject, Subject 6, scored within a 10 percent range for all three tasks. The greatest scoring diversity was exhibited by Subject 7, whose scores ranged from 21.3 in conversation to 62.5 percent in monologue. Like the Mean of Judges results, scores recorded by the Experimenter also varied greatly. The three total means for disfluent words were very similar for the three speaking tasks. The three means of the speaking tasks ranged within a 2.9 percent spread. These were reading 44.5, monologue 47.3, and conversation 44.4 percent. The means of the percent of time speaking were all above 97.2 percent.

In comparing Table 3, Mean of Judges, to Table 4, Experimenter, there is a very close relationship between the total means recorded for each speaking task and speaking time. In the reading task the mean of Judges recorded a 46.3 percent disfluency and a 100 percent speaking time. For the same task the Experimenter reported a 44.5 percent

disfluency and a 100 percent speaking time. In the monologue task the Mean of Judges listed a 46.3 percent disfluency mean and a 99.4 percent speaking time. The Experimenter recorded a 47.3 percent disfluency mean and a 99.1 percent speaking time. In conversation the disfluency mean was 47.8 percent for the Mean of the Judges and a 44.4 percent for the Experimenter. The mean for speaking time was 97.7 percent for the Mean of the Judges and 97.2 percent for the Experimenter. These results show that the Mean of the Judges' data and the Experimenter's data for total mean disfluency and speaking time, are within 3.0 percent on every speaking task. However, the wide diversity of scores per subject recorded by both the Experimenter and the Mean of the Judges demonstrates the difficulty in predicting one score on the basis of another for the three speaking tasks.

Coefficients of Correlation Between Sander's Index and TIDT

Table 5 shows the coefficients of correlation of percent disfluent words between Sander's speaking tasks and the TIDT speaking tasks for the Mean of the Judges and the Experimenter. In comparing the TIDT reading task to Sander's reading task the Mean of the Judges shows a coefficient of correlation of .82 and as does the Experimenter. This coefficient of correlation comparing TIDT reading to Sander's reading is significant at the .01 level of confidence. Comparing the TIDT reading task to Sander's Job Task the Mean of the Judges records a coefficient of correlation of .82, which is significant at the .01 level of confidence. The Experimenter's data comparing the TIDT reading task to the Job Task yields a coefficient of correlation of .71, which is significant at the .05 level of confidence. When TIDT

TABLE 5. Coefficients of correlation of percent disfluent words between Sander's speaking tasks and the TIDT speaking tasks for the Mean of the eleven Judges and the Experimenter.

MEAN OF ELEVEN JUDGES

	SANDER	
TIDT	Reading	Job Task
Reading	.82**	.82**
Monologue	.29	.49
Conversation	.39	.63

EXPERIMENTER

	SANDER	
TIDT	Reading	Job Task
Reading	.82**	.71*
Monologue	.46	.49
Conversation	.36	.84**

* Significant at .05 level of confidence
** Significant at .01 level of confidence

monologue is correlated to Sander's reading task the coefficient of correlation is .29 for the Mean of the Judges, while the Experimenter's coefficient of correlation is .46. The TIDT monologue task compared with Sander's Job Task yields a coefficient of correlation of .49 for the Mean of the Judges and the Experimenter. The TIDT monologue correlation to both the Sander's reading and Job Task does not show a significant relationship for either the Mean of Judges' analysis or the Experimenter's analysis. Looking at TIDT's conversation task correlated to Sander's reading task, the coefficient of correlation is .39 for the Mean of the Judges and .36 for the Experimenter. Neither of these scores are significant. Correlating the TIDT's conversation to Sander's Job Task yields a coefficient of correlation of .63 for the Mean of Judges and .84 for the Experimenter. The Mean of the Judges coefficient of correlation of .64 is close to but below the .05 level of confidence for this task. However, the Experimenter's coefficient of correlation for this task, TIDT conversation vs. Sander's Job Task, shows a .84 which is significant at the .01 level of confidence.

Table 6 shows coefficients of correlation of percent speaking time between Sander's speaking tasks and the TIDT speaking tasks for the Mean of the Judges and the Experimenter. No coefficients could be computed correlating the percent speaking time of the TIDT to the speaking task to either Sander's reading task or Job Task speaking time. This was shown in both the Mean of the Judges and the Experimenter's results. In comparing the TIDT monologue to Sander's reading task a coefficient of -.64 was obtained by the Mean of the Judges and the Experimenter. This coefficient is slightly below the .05 level of confidence. The TIDT monologue task correlated to Sander's Job Task

TABLE 6. Coefficients of correlation of percent speaking time (rate) between Sander's speaking tasks and the TIDT speaking tasks for the Mean of the eleven Judges and the Experimenter.

MEAN OF ELEVEN JUDGES

	SANDER	
TIDT	Reading	Job Task
Monologue	-.64	-.81**
Conversation	-.06	-.17

EXPERIMENTER

	SANDER	
TIDT	Reading	Job Task
Monologue	-.64	-.80**
Conversation	-.20	-.16

** Significant at .01 level of confidence

yields a coefficient of correlation of -.81 for the Mean of the Judges and -.80 for the Experimenter. Both the Mean of the Judges and the Experimenter's coefficients of correlation are significant at the .01 level of confidence. The TIDT conversation task correlated to Sander's reading task shows a coefficient of correlation of -.06 for the Mean of Judges and -.20 for the Experimenter's data. The TIDT conversation task correlated to Sander's Job Task shows a coefficient of correlation of -.17 for the Mean of the Judges and -.16 for the Experimenter. For both the Experimenter and the Mean of the Judges' data the TIDT conversation task correlated to Sander's reading or Job Task does not yield a coefficient of correlation that is significant. The negative sign recorded on all the coefficients of correlation for comparing rate is caused by the fact that there is an inverse relationship between Sander's measurement of speaking time and the TIDT's measurement of speaking time. The Sander's measurement for speaking time involves counting total number of seconds to finish a task. The TIDT involves recording the percent of time speaking compared to the total time available. Therefore, for Sander's the more fluent a subject the less number of seconds recorded but for the TIDT the more fluent a subject a higher percentage of speaking time recorded.

CHAPTER IV

DISCUSSION

The results of this research, studying the interrelationships between Sander's Disfluent Word Index and the TIDT, provided several comparisons between the two procedures. Based on the significant correlations obtained between the TIDT reading task and Sander's reading and Job Task, it can be said that the TIDT reading task is an effective predictor of both the Sander's reading and Job Task. This is shown in both the Mean of the Judges and the Experimenter's data. Furthermore, there is a possibility that the TIDT conversation task is a predictor of Sander's Job Task. The Mean of the Judges' data for this correlation is slightly below the significant level at the .05 level of confidence. However, the Experimenter's results show a significant correlation between the TIDT conversation and Sander's Job Task. This may possibly be explained by the instructions of the Job Task and what is required to complete it.

Looking at the Job Task carefully it should be noted it is not the same as a pure monologue situation but has qualities which make it a conversational task also. The directions involved in obtaining a Job Task sample require the subject to tell about his present job, past jobs, or future jobs. Since usually this is a one to one situation with only the subject and the examiner in the room it demonstrates some qualities of a one sided dialogue. The subject has someone to talk to, even though the examiner is silent. The subject is required to organize his thoughts and present them on a topic he has probably discussed

with many people before. With the one to one relationship, talking about a familiar topic, and having the benefit of the examiner's presence and silent acknowledgement, it can be understood why all these factors make it more of a conversational task than strict monologue. Furthermore, the directions for the Job Task state that if the subject does not talk long enough for a three minute or 250 word sample, the examiner is required to prompt him verbally to continue speaking. When you add this factor to the others involved in the Job Task it takes on even more qualities of a conversation.

This research also showed the lack of a significant correlation between the TIDT monologue and the Job Task. Since both are considered monologues this lack of a significant correlation is surprising. However, if indeed the Job Task is more of a conversational task than a monologue this insignificant correlation is justified. In fact, the low correlation between the TIDT monologue and the Job Task suggests that the monologue is measuring another type of behavior. In contrast to the Job Task, the TIDT monologue had quite different characteristics. This task required the subject to look at pictures, one at a time, and talk on any part or detail of the picture he wished. In the Job Task where the subject faced the examiner and talked "to" him, the TIDT monologue required the subject to pay attention only to the various pictures, not making him feel the necessity to relate to another person on a one to one basis. He was merely required to think aloud what he saw or came to his mind while studying the picture. The presence of another person in the room was forgotten. Also, the listener was not to prompt in any way and with the vast amount of pictures presented to the subject there was never a need to prompt. The coefficient of correlations reported

in this study support the conclusion that indeed the TIDT measures not only the behaviors that Sander's tasks measure but much more than Sander's speaking tasks.

Looking strictly at the TIDT scores it should be stated that these mean scores for each task point out that indeed they are separate entities measuring three separate sets of behavior. The mean scores for each task are so close, within a three percent range, that it is impossible to predict a mean score of a speaking task on the basis of another task. Looking at individual scores of a subject for the three speaking tasks serves to emphasize this point. The three mean scores for any subject are so diverse that a trend for predicting scores is impossible to establish. This is shown in both the Mean of the Judges scores and the scores of the Experimenter. This reiterates the point that using the three TIDT speaking tasks is necessary and provides more information than Sander's two speaking tasks.

Other important factors not shown in the data analysis concern the time and preparation necessary to analyze this data using Sander's procedures and the TIDT procedures. Besides using these two procedures to analyze the data so correlations could be made, another important purpose was to compare the two procedures strictly in terms of administrating the two procedures. The Sander's procedure required the judges to have a prepared transcript for each speaking task. Time was needed to record the total number of seconds required to perform each task, making sure any time used by the examiner in speaking or prompting was deleted from the subject's total time. After this preparation was completed, each judge was to listen to each speaking sample three times to insure accuracy in identifying disfluencies. The judge was required to identify

each disfluency while following along with the transcript and then tabulate the total disfluency score. This preparation and analysis required 2-4 hours per subject. The TIDT scoring procedure required only a score sheet and three two hour sessions to complete all nine subjects. The nature of the scoring, marking whether a disfluent word had occurred within a 10 second interval, requires no pre-scoring preparation such as a transcript and timing each sample. The TIDT speaking sample is exactly four minutes long and scoring of disfluency and percent talking is done simultaneously instead of being two separate tasks. The difference in time spent analyzing the samples was six hours for nine subjects or 15 minutes per subject using the TIDT and 2-4 hours for one subject using Sander's procedure. This is magnified even more by the fact that the TIDT had one more task than the Sander's procedure.

Another factor that should be mentioned when comparing the TIDT to Sander's procedure, involves the administration of the test rather than the analysis of the sample. The examiner in the room with the subjects found it difficult to obtain a 250 word sample of the Job Task. The Job Task is supposed to be a three minute sample; however, Sander felt 250 words were necessary to obtain an effective speech sample. The subject seemed to have difficulty saying 250 words about his past, present, or future job. A third of the nine subjects had to be verbally prompted and some others visibly struggled with what to say next. Often the Job Task ran over five minutes with long pauses by the subject and verbal prompts by the examiner to insure a 250 word sample. This problem in administrating the Job Task plus the obvious time factor in analyzing Sander's speaking tasks make the TIDT a much more effective and quicker test to administer and score.

There are some aspects of this research that point out weaknesses that need to be discussed. The most obvious weakness shows itself in the coefficients of correlations between Sander's speaking tasks and the TIDT speaking tasks in terms of speaking time (rate). In the Experimenter's and Mean of the Judges' data no coefficient of correlation could be computed for the TIDT reading compared to Sander's reading and Job Task. This is probably caused by the fact that in the TIDT reading task the percent speaking time was always 100 percent. There was no range between the subjects for this task. Because it was a reading task the subject was speaking during every 10 second interval so a score of 100 percent speaking time was recorded. Another problem arose in that only significant correlation of speaking time was between the TIDT monologue and Sander's reading and Job Task. This was true for both the Mean of the Judges and the Experimenter's data. This is difficult to explain because the percent disfluent words correlated poorly when comparing TIDT monologue to Sander's reading and Job Task. Yet, the speaking time demonstrated a high significant correlation. The TIDT conversation correlated to Sander's reading and Job Task for percent speaking time showed a low coefficient of correlation when it was expected that it would correlate better than the TIDT monologue task. Part of the problem may be in the lack of range for percent speaking time of the TIDT. The most diverse range was the TIDT conversation task but this was less than a three percent range. A possible improvement may be to lower the 10 second interval to five seconds. However, time samples of less than 10 seconds are difficult to do by hand and would probably require a mechanical apparatus to insure accuracy. The question arises whether a measure of percent speaking time is necessary at all. For the TIDT the

need to measure percent speaking time is a critical factor. Without some knowledge of what percent of the time available a subject was speaking it would be impossible to know how significant a percent disfluent score would be. If you have a subject who was 25 percent disfluent in a four minute block how reliable would that percent be without some indication of what percent he had spoken within that four minute sample. For the TIDT a percent speaking time is necessary and an effective measure to correspond with the percent disfluent score. However, this measurement of percent speaking time does not correlate significantly with Sander's measure of speaking time. The complete reason why is still unclear.

Another problem which had a direct effect upon this research revolves around the comparison of ranges of subjects for Sander's (1961) study and the present research. The obvious difference is number of subjects. Sander had 40 subjects and this research had nine subjects. Also, the subjects Sander had demonstrated more stuttering behavior so his range of disfluent words far exceeds the range of this research. The minimum scores for Sander and this research are very similar in both percent disfluent words and speaking time. However, Sander's maximum scores extend far beyond the maximum scores for the present research subjects. This is a result of having three times as many subjects plus having more severe stutterers. The fact that these two groups of subjects compared to each other are so different in number and severity makes a direct comparison very difficult. Possibly with more subjects this research would have been better able to replicate Sander's study in terms of number and severity of subjects. Having more diversity in subjects and severity would have strengthened the findings within this study.

This research provided information on a diagnostic tool, the TIDT. Comparing the interrelationships of the TIDT to Sander's Disfluent Word Index provided a comparison of a relatively new disfluency measure to an established method of disfluency analysis. On the basis of the facts related about Sander's Index and the TIDT, it is speculated that while the TIDT reading task can predict the Sander's reading and Job Task, the monologue and conversation tasks measure even more behaviors than the Sander's Job Task. Therefore, the TIDT provides more information about the subject's behavior in a variety of situations. The two Sander's speaking tasks do not provide the amount or diversity of information acquired by the TIDT. The results of the TIDT scores show each task cannot be predicted from another task. Each task is a separate entity, each with a place in the diagnostic evaluation. Consequently, upon studying the administration and analysis of these two procedures, Sander's Disfluent Word Index and the TIDT, shows the TIDT to be valid, efficient, and overwhelmingly faster and easier to utilize in assessing disfluency.

Implications for Application and Further Research

There is a need to investigate the possibilities of using a five second interval instead of the 10 second interval utilized in the TIDT. The five second interval could improve the measuring of percent speaking time in the TIDT. A 10 second interval seems to be too long because of the lack of range shown in the group of subjects in this research. The range for the mean of the three tasks in percent speaking time was less than three percent. A study could be done to see if a five second interval provides more accurate information required for

the TIDT. Since manually it would be difficult to record disfluencies within a five second interval, a mechanical apparatus would have to be utilized to insure accuracy. Perhaps the five second interval would solve the problems encountered in this research in regard to measuring speaking time (rate).

Along this same line more research on the TIDT might be suggested to examine more closely the TIDT's value as a therapeutic tool. This research has concerned itself with the diagnostic value of the TIDT in comparison to another diagnostic tool, Sander's Disfluent Word Index. The possible therapeutic uses should be examined and compared with other tools used in the therapy situation. The author feels the TIDT has a place in the therapy session to assess progress in the therapy situation. The ease and quickness of the TIDT make it amenable to a therapy session. The potential of the TIDT for diagnostic purposes is evident but its therapeutic values should be explored to examine all possibilities.

CHAPTER V

SUMMARY

The purpose of this research was to study the interrelationships between Sander's Disfluent Word Index and the Time-Interval Disfluency Test. Both tests are diagnostic tools used to assess disfluencies and speaking time in the evaluation of stuttering.

The Sander's Disfluent Word Index consists of two speaking tasks, reading and Job Task (monologue). In the reading task the subject has to read a 300 word passage. In the Job Task the subject is instructed to tell about past, present, or future jobs. To score these two tasks the total number of disfluent words identified in each sample are divided by the total number of words for each speaking task yielding a disfluency percentage. The speaking time is obtained by recording the total number of seconds required to complete each of the two speaking tasks.

The TIDT consists of three speaking tasks, reading, monologue and conversation. The reading task consists of cards containing a different exerpt from the encyclopedia printed on each. The monologue task consists of cards containing a photograph from various magazines. The conversation task involves the examiner asking the subject various questions to elicit a typical conversation. Each speaking task is exactly four minutes long. Recorded simultaneously with the speech sample are beeps every 10 seconds. The beeps are used to sub-divide the sample into 10 second intervals. In scoring, the listener marks on the score sheet only if a disfluency occured during the 10 second

interval, not how many disfluencies occured. To obtain a percent disfluent score the number of intervals where a disfluency occured are divided by the total number of intervals where the subject was speaking. The percent speaking time is scored by counting the total number of 10 second intervals where the subject was speaking and dividing it by the total number of intervals available.

Nine male stutterers served as subjects. The two Sander's tasks and the three TIDT speaking tasks were presented to the subjects each in a different order. Scoring of the samples was done by eleven judges who were trained to use both Sander's Index and the TIDT. The data analysis consisted of coefficients of correlations obtained comparing each of the Sander's speaking tasks to the TIDT speaking tasks. The percent speaking time for Sander's tasks and the TIDT tasks were also correlated.

The results of the research indicate the TIDT reading task shows a significant coefficient of correlation to both the Sander's reading and Job Task. The TIDT monologue task did not show a significant relationship to either Sander's reading or Job Task. The TIDT conversation task showed a significant coefficient of correlation when compared to Sander's Job Task but not the reading task. The TIDT scores showed a wide range of scores for each speaking task. However, the mean scores for each speaking task ranged less than three percent. When comparing the administration and scoring for the two procedures, the TIDT takes 15 minutes to score one subject where the Sander's Index takes 2-4 hours to score one subject.

These correlations indicate that indeed the TIDT is measuring the same behaviors as the Sander's Index. Moreover, the TIDT appears

to be measuring more behaviors than Sander's Index. Looking at the TIDT scores it is shown that on the basis of the three mean scores for the three tasks each is a separate entity. The three mean scores are so similar it would be impossible to predict one score on the basis of another. In terms of administration the TIDT shows itself to be faster and easier to score than Sander's Index. Consequently, the TIDT shows itself to be valid, efficient, and more expedient than the Sander's Disfluent Word Index.

APPENDICES

APPENDIX A

INSTRUCTIONS GIVEN TO SUBJECT BY EXAMINER FOR SANDER'S READING PASSAGE AND TIDT'S READING CARDS

Sander's Reading Passage: Please read this passage aloud. Just read it as you normally would. Continue until you are asked to stop.

TIDT Reading Cards: Here are a group of cards. Each card contains a different paragraph. Take each card and read the complete paragraph as you normally would. When you've finished with a card go on to the next card. Continue reading until you are asked to stop.

APPENDIX B

INSTRUCTIONS GIVEN TO SUBJECT BY EXAMINER FOR JOB TASK AND TIDT PICTURE TASK (MONOLOGUE)

Job Task: Now I'm going to ask you to tell me all about your future occupation. Describe it in as much detail as possible. I will tell you when to stop. If you run out of words before I tell you to stop go on and describe jobs you have previously held or a future job you would like to hold. Please continue speaking until I say stop. Take a minute to organize your thoughts then tell me when you're ready to begin talking.

TIDT Picture Task (Monologue): Here are a group of cards with pictures on them. Take each picture one at a time and describe what is happening in it or make any comment about the picture you wish. It's not important what you say about it, only that you continue talking. After you've said all you can about the picture go on to the next one and do the same thing. Continue through the group of pictures until asked to stop.

APPENDIX C

QUESTIONS USED IN TIDT CONVERSATION TASK

1. What types of leisure activities do you engage in?
2. Have you ever participated in any organizations?
3. How active are you in political affairs?
4. How would you describe your personality characteristics?
5. Where are you living on campus? Do you like it?
6. How relevant do you feel your education was for you?
7. Do you feel that coming to K.S.U. (K.U.) has affected your attitudes in any way?
8. What would you have liked to change about your schooling so far?
9. What field of study are you in?
10. After you obtain your degree are you planning on graduate study?
11. Do you have any brothers and sisters? Can you tell me about them?
12. How would you describe your relationship with your parents?
13. What places have you traveled to?
14. Where would you like to go if you had the opportunity? Why?
15. What types of things really irritate you about people or situations?
16. What particular traits do you admire in people?
17. If there was one thing you could change about yourself what would it be?
18. What do you feel your best qualities are?
19. What person have you most admired? Why?

APPENDIX D

INSTRUCTIONS TO ELEVEN JUDGES FOR SANDER'S ANALYSIS OF SPEECH SAMPLES (HAND-OUT)

Definition of a Disfluent Word: A word was considered to be disfluent if it involved prolonged sounds, was classified as a broken word, was involved in a sound, syllable or word repetition, or was interrupted by an interjection. Words preceeded by interjections or involved in phrase repetitions were not counted as disfluent words.

Instructions: Two speaking tasks will be involved. The first is an oral reading passage. The second is a monologue of the Job Task. Listen to each speaking situation 3 times to insure accuracy. Tabulate the total number of disfluent words for each speaking task and record the number at the bottom of the transcript for each task.

Please return your data sheets for each subject to Room 5. Please put your initials on the data sheet of each completed subject you return.

STUTTERING ANALYSIS
DUE 3-9-73

REFERENCES

Ayllon, T. and Azrin, N. H., The Token Economy: A Motivational System for Therapy and Rehabilitation. New York: Appleton-Century-Crofts, Inc., 1968.

Berry, R. and Silverman, F., Equality of Intervals on the Lewis-Sherman Scale of Stuttering Severity. Journal of Speech and Hearing Disorders, 15, 1972, 185-188.

Bijou, W., Peterson, R. F., and Ault, M. H., A Method of Integrating Descriptive and Experimental Field Study at the Level of Data and Empirical Concepts. Journal of Applied Behavior Analysis, 1, 1968, 175-191.

Buetzer, C. B., Reliability of the Time-Interval Disfluency Test on an Inter-observer Agreement Basis. Kansas State University, 1973.

Cullinan, W. L., Prather, E. M., and Williams, D. E., Comparison of Procedures for Scaling Severity of Stuttering. Journal of Speech and Hearing Research, 1, 1963, 187-193.

Duffy, R. J., A Quantitative Study of the Speech Nonfluencies of Fifty Adult Female Stutterers. A Master's Thesis, University of Iowa, 1957.

Fairbanks, G., Voice and Articulation Drillbook. New York: Harper & Brothers, 1940.

Johnson, W., Measurements of Oral Reading and Speaking Rate and Disfluency of College Age Male and Female Stutterers and Non-Stutterers. Journal of Speech and Hearing Disorders, Monograph Supplement 7, 1961, 1-20.

Johnson, W., Darley, F. L. and Spriesterbach, D. C., Diagnostic Methods in Speech Pathology. New York: Harper & Row, 1963.

Johnson, M., Kelley, S., Harris, F. and Wolf, M. M., An Application of Reinforcement Principles to Development of Motor Skills. Child Development, 37, 1966, 379-387.

Lewis, D. and Sherman, D., Measuring the Severity of Stuttering. Journal of Speech and Hearing Disorders, 16, 1951, 320-326.

Riley, G. D., A Stuttering Severity Instrument for Children and Adults. Journal of Speech and Hearing Disorders, 37, 1972, 314-321.

Sander, E. K., Reliability of Iowa Speech Disfluency Test, Journal of Speech and Hearing Disorders, Monograph Supplement 7, 1961, 21-30.

Starbuck, H. B., Determination of Severity of Stuttering and Construction of an Audio-Visual Scale, Doctoral Dissertation, Purdue University, 1954.

Van Riper, C., The Nature of Stuttering. Prentice-Hall, Inc., Englewood Cliffs, N. J., 1971.

Wertheim, E., A New Approach to Classification and Measurement of Stuttering, Journal of Speech and Hearing Disorders, 37, 1972, 242-251.

Young, M. A., Predicting Ratings of Severity of Stuttering, Journal of Speech and Hearing Disorders, Monograph Supplement 7, 1961, 31-54.

_____ Anchoring and Sequence for the Category Scaling of Stuttering Severity, Journal of Speech and Hearing Research, 13, 1970, 360-368.

A STUDY OF THE INTERRELATIONSHIPS BETWEEN THE TIME-INTERVAL DISFLUENCY TEST AND SANDER'S DISFLUENT WORD INDEX

by

CAROL COBURN HAMMOND

B. A., Kansas State University, 1972

AN ABSTRACT OF A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF ARTS

Department of Speech

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1973

The purpose of this research was to study the interrelationships between Sander's Disfluent Word Index and the Time-Interval Disfluency Test. Both tests are diagnostic tools used to assess disfluencies and speaking time in the evaluation of stuttering.

The Sander's Disfluent Word Index consists of two speaking tasks, reading and Job Task (monologue). In the reading task the subject has to read a 300 word passage. In the Job Task the subject is instructed to tell about past, present, or future jobs. To score these two tasks the total number of disfluent words identified in each sample are divided by the total number of words for each speaking task yielding a disfluency percentage. The speaking time is obtained by recording the total number of seconds required to complete each of the two speaking tasks.

The TIDT consists of three speaking tasks, reading, monologue and conversation. The reading task consists of cards containing a different exerpt from the encyclopedia printed on each. The monologue task consists of cards containing a photograph from various magazines. The conversation task involves the examiner asking the subject various questions to elicit a typical conversation. Each speaking task is exactly four minutes long. Recorded simultaneously with the speech sample are beeps every 10 seconds. The beeps are used to sub-divide the sample into 10 second intervals. In scoring, the listener marks on the score sheet only if a disfluency occured during the 10 second interval, not how many disfluencies occured. To obtain a percent disfluent score the number of intervals where a disfluency occured are divided by the total number of intervals where the subject was speaking.

The percent speaking time is scored by counting the total number of 10 second intervals where the subject was speaking and dividing it by the total number of intervals available.

Nine male stutterers served as subjects. The two Sander's tasks and the three TIDT speaking tasks were presented to the subjects each in a different order. Scoring of the samples was done by eleven judges who were trained to use both Sander's Index and the TIDT. The data analysis consisted of coefficients of correlations obtained comparing each of the Sander's speaking tasks to the TIDT speaking tasks. The percent speaking time for Sander's tasks and the TIDT tasks were also correlated.

The results of the research indicate the TIDT reading task shows a significant coefficient of correlation to both the Sander's reading and Job Task. The TIDT monologue task did not show a significant relationship to either Sander's reading or Job Task. The TIDT conversation task showed a significant coefficient of correlation when compared to Sander's Job Task but not the reading task. The TIDT scores showed a wide range of scores for each speaking task. However, the mean scores for each speaking task ranged less than three percent. When comparing the administration and scoring for the two procedures, the TIDT takes 15 minutes to score one subject where the Sander's Index takes 2-4 hours to score one subject.

These correlations indicate that indeed the TIDT is measuring the same behaviors as the Sander's Index. Moreover, the TIDT appears to be measuring more behaviors than Sander's Index. Looking at the TIDT scores it is shown that on the basis of the three mean scores for the three tasks each is a separate entity. The three mean scores are

so similar it would be impossible to predict one score on the basis of another. In terms of administration the TIDT shows itself to be faster and easier to score than Sander's Index. Consequently, the TIDT shows itself to be valid, efficient, and more expedient than the Sander's Disfluent Word Index.