Validity and Efficiency of the Check-Slash Transcription Method for Measuring Intelligibility

Vicky Jo Bacon
Portland State University

Recommended Citation

This Thesis is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. For more information, please contact pdxscholar@pdx.edu.
THESIS APPROVAL

The abstract and thesis of Vicky Jo Bacon for the Master of Science in Speech Communication: Speech and Hearing Sciences were presented May 10, 1995, and accepted by the thesis committee and the department.

COMMITTEE APPROVALS: 

Mary Gordon-Branman, Ph.D., Chair

John A. Tetnowski, Ph.D.

Jeanette S. DeCarrico, Ph.D.
Representative of the Office of Graduate Studies

DEPARTMENT APPROVAL: 

Stephen A. Kosokoff, Chair
Department of Speech Communication

*********************************************************

ACCEPTED FOR PORTLAND STATE UNIVERSITY BY THE LIBRARY

by __________________________ on 29 June 1995
ABSTRACT


Title: Validity and Efficiency of the Check-Slash Transcription Method for Measuring Intelligibility

Speech-language pathologists are routinely called upon to make professional assessments concerning a speaker's level of intelligibility. The use of subjective judgement procedures for estimating a percentage of intelligibility is the general practice of many speech-language pathologists because they require minimal time. Although efficient, these methods lack any form of numerical support, and their validity and reliability is questionable. The standard within the field that provides data support is the orthographic transcription method, but it is considered to be too time-consuming for practical application (Samar & Metz, 1988). Researchers continue to seek a measure that is both valid and efficient to be used clinically.

The purpose of this study was to establish validity of a check-slash transcription method used to provide objective numerical support for assigning percentage of intelligibility for individuals with moderate speech
impairments. The study sought to answer the following questions:

1) Is the check-slash method of transcription a valid measure for quantifying percentage of intelligibility?
2) Is the check-slash method a more time-efficient procedure than the orthographic transcription method?

The subjects for this study were 20 graduate students from Portland State University, that were randomly assigned to two transcription groups (check-slash or orthographic). Each listener transcribed 12 samples taken from 2 girls and 10 boys between the ages of 4:1 and 5:6 with a moderate degree of phonological deficiency.

The data were analyzed using individual Mann-Whitney U Tests for each of the 12 samples. Results indicated no significant difference between the check-slash and orthographic transcription methods when used to assign a percentage of intelligibility to individuals with a moderate speech deficit. Although no significant difference was found, interrater reliability for both methods was low.

This study established efficiency for the check-slash transcription method when compared to the orthographic method. Increased efficiency for the check-slash method ranged from 38% to 54% over the orthographic method. Results may also indicate that listener perception may influence each clinician's ability to be accurate in their assessments.
VALIDITY AND EFFICIENCY OF THE CHECK-SLASH TRANSCRIPTION METHOD FOR MEASURING INTELLIGIBILITY

by

VICKY JO BACON

A thesis submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE
in SPEECH COMMUNICATION: SPEECH AND HEARING SCIENCES

Portland State University
1995
ACKNOWLEDGEMENTS

This sincere word of thanks goes to my husband Doug, for his unending support and words of encouragement. Without his love I never could have accomplished what I have. A special thank you for my daughters Angela and Jenifer. I hope I have instilled in you the importance of education, and the desire to reach for the stars.

I also wish to thank my two best friends, Deborah and Debi, for listening to me whine and complain all along the way. Finally, a heartfelt thank you to all of the Speech and Hearing Sciences faculty at Portland State University for giving me the knowledge I needed to fulfill my dream.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................... ii
LIST OF TABLES ................................................. v

CHAPTER

I  INTRODUCTION AND STATEMENT OF PURPOSE ........... 1
   Introduction .................................................. 1
   Statement of Purpose ...................................... 4

II  REVIEW OF THE LITERATURE ............................... 6
   Speech Intelligibility Defined ......................... 6
   Measures of Intelligibility ............................ 7
      Subjective Measures of Intelligibility
      Objective Measures of Intelligibility
   Summary of Research Needs ............................ 17

III  METHODS AND PROCEDURES .............................. 19
   Subjects ..................................................... 19
   Measuring Instruments ................................... 20
   Procedures .................................................. 20
      Speech Samples
      Transcription Methods
      Transcription Setting
      Transcription Procedures

   Scoring and Data Analysis .............................. 25
      Reliability
RESULTS AND DISCUSSION ......................................................... 27

Results ................................................................................. 27

Reliability Analysis
Intelligibility Analysis
Time Analysis

Discussion ................................................................. 33

Intelligibility
Time

SUMMARY AND IMPLICATIONS ......................................................... 43

Summary ................................................................. 43

Implications ............................................................... 45

Clinical
Research

REFERENCES ................................................................. 48

APPENDICES

A FACTORS THAT AFFECT INTELLIGIBILITY ........... 53

B TRANSCRIPTION FORMS ......................................................... 55

C TRANSCRIPTION INSTRUCTIONS ......................................................... 58

D RAW DATA ................................................................. 63
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Comparison of Judges by Pairs for the Check-Slash Method of Transcription across the 12 Samples</td>
</tr>
<tr>
<td>2</td>
<td>Comparison of Judges by Pairs for the Orthographic Method of Transcription across the 12 Samples</td>
</tr>
<tr>
<td>3</td>
<td>Descriptive Statistics for Percentage of Intelligibility. Comparison of the Check-Slash and Orthographic Methods of Transcription</td>
</tr>
<tr>
<td>4</td>
<td>Descriptive Statistics: Transcription Time Required to Complete each Sample using Both Methods</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION AND STATEMENT OF PURPOSE

Introduction

Successful interactions with others are dependent on many factors. One of utmost importance is the ability of the listener to understand the words of the speaker. This is what is referred to as speech intelligibility. Speech-language pathologists are routinely called upon to make professional judgements concerning a speaker's level of intelligibility. These assessments are necessary to quantify the severity of disordered communication. The numerical result, stated in a percentage form, is often used to qualify a patient for services; document treatment need and progress; convey understandable information to patients, caregivers, and other professionals; and reflect the speaker's functional communication ability.

Individual clinicians have their own ways of assessing intelligibility, as there are many different methods available. One approach is administering standardized tests that quantify a client's level of intelligibility in either isolated words or contextual speech (e.g., Weiss, 1982; Yorkston & Beukelman, 1981). Many non-standardized measures are also used, including evaluating the client's
intelligibility using a rating scale, gross estimations of percentage of intelligibility based on perceptual judgements, and collecting a language sample and transcribing it as the basis for determining the percentage of intelligible words. No matter which approach is chosen, three factors are of key importance: reliability, validity, and efficiency.

As medical costs soar, efficacy of treatment has become a major issue in the medical field, and services provided by speech-language pathologists are no exception. It is necessary to document patient progress, not only for insurance providers, but for patients and caregivers as well. Proper documentation provides support for both treatment need and progress. For this reason, the method chosen needs to be a reliable and valid measure. Subjective measures often fail to meet this criterion.

Due to increasingly heavy caseloads of practicing speech-language pathologists, efficiency has now become an important factor. For years, the orthographic or write-down method of transcription (word-for-word recording) has been considered the traditional method. As a measure of intelligibility, it has a high degree of face validity (Kwiatkowski & Shriberg, 1992; Samar & Metz, 1988). Many speech-language pathologists, however, agree it often becomes too time-consuming and other methods are chosen. Currently the general practice of many speech-language
pathologists is to estimate the percentage of intelligibility of a client (Gordon-Brannan, 1993, 1994; Kent, Milolo, & Blodel, 1994; Samar & Metz, 1988; Shriberg & Kwiatkowski, 1982; Weiss, 1982; Yorkston & Beukelman, 1978). Such estimates may be neither reliable nor valid.

In 1982, Weiss, believing that standardization was necessary for quantifying intelligibility, created the Weiss Intelligibility Test. Within this test, he created an alternative to orthographic transcription for the assessment of contextual speech, which consists of a recording grid and two symbols, a check (√) to represent intelligible words and a slash (−) to denote unintelligible words. This will hereafter be referred to as the check-slash method.

The check-slash method of transcription is unlike the orthographic method, as the clinician does not write down every word the client has said; rather the words are recorded by using the appropriate symbol, as either being understood (intelligible) or not understood (unintelligible). This method has the probability of being less time-consuming. The question then becomes: Is the check-slash method as valid as the orthographic method in the assessment of percentage of intelligibility of contextual speech? If the check-slash method can be shown to yield data similar to the results of the orthographic method and it is shown to be more time efficient, then this procedure would be considered a useful clinical tool. A
method such as this would encourage clinicians to move away from the subjective measures of gross estimation and scaling procedures and move toward a more precise way of quantifying speech intelligibility supported by numerical data.

Statement of Purpose

The purpose of this study was to determine the validity and efficiency of the check-slash method of transcription for percentage of intelligibility as compared to the orthographic method when assessing speakers with moderate phonological deficiencies. The research questions addressed were:

1. Is the check-slash method of transcription a valid measure for quantifying percentage of intelligibility, when used to assess children with moderate phonological deficits?

2. Is the check-slash method of transcription a more time-efficient procedure for assessment of children with moderate phonological deficits, when compared to orthographic transcription?

These research questions are reflected in the two null hypotheses for this study:

1. There is no significant difference between the results obtained using the check-slash transcription method as compared to the orthographic method of transcription when used to compute the percentage of intelligibility for
children with moderate phonological deficits.

2. There is no significant difference in the average amount of time necessary to transcribe a language sample using either the check-slash or orthographic methods of transcription, when computing a percentage of intelligibility index, for children with moderate phonological deficits.
CHAPTER II

REVIEW OF THE LITERATURE

In this literature review the focus is on two aspects of intelligibility: definition and most commonly used methods of measurement. Much of the literature to date has been focused on intelligibility measurements of speakers with dysarthria, hearing impairment, and laryngectomy.

Intelligibility Defined

Gordon-Brannan (1993) defined speech intelligibility as "the degree to which a person's speech is understood by a listener" (p. 7). This degree of intelligibility is often expressed as a percentage of words understood. A person's intelligibility, therefore, can range from being totally understood (100% intelligible) to not being understood at all (0% intelligible) and all the levels in between. Kent et al. (1994) referred to intelligibility as the "functional common denominator of verbal behavior" (p. 81). Based on these definitions, it is therefore understandable that, without a reasonable amount of intelligibility, a person's ability to communicate with others will be limited. There is no doubt that intelligibility is difficult to measure, as it can be influenced by so many contributing factors. In
his intelligibility test manual, Weiss (1982) identified 22 factors that can affect an individual's ability to be understood (Appendix A). His list, however, did not include phonological deviations used by children that can also lead to unintelligibility (Hodson & Paden, 1981). Another aspect contributing to this difficulty is the subjective nature of the task and its dependence on listener perception.

Measures of Intelligibility

In the literature, a wide variety of methods to assess the intelligibility levels of clients with communication disorders have been described. Some procedures are clearly very subjective in nature, whereas others are more objective. As new objective measures become available such as the Preschool Speech Intelligibility Measure (P-SIM) (Wilcox, Schooling, & Morris, 1991), other subjective measures, such as equal-interval scaling, have been criticized for lack of criterion validity (Samar & Metz, 1988). The issue of quantifying intelligibility levels with objective numerical data, rather than subjective measures, has become more prevalent as the current trend in the field of speech-language pathology moves toward efficacy of treatment, and accountability.

The literature review conducted by this researcher did not reveal information relating directly to the Weiss (1982)
check-slash method of transcription for contextual speech. This could be because this transcription method was not designed to be used alone, but in conjunction with a single word identification task. This is only one method of measuring intelligibility. Some of the more clinically popular methods are reviewed below.

Subjective Measures of Intelligibility

Subjective measures are based on the perceptual judgements of the listener. Using rating scales in a variety of forms, the clinician assigns a numerical score to the intelligibility of the speaker.

Scaling Procedures. Scaling procedures have long been used clinically and in research related to speech intelligibility and communication disorders. Two types of scaling procedures are equal-interval rating and direct magnitude estimation.

The most commonly used interval scaling technique for rating intelligibility is the equal-appearing interval scale (Schiavetti, 1992). In this method, intelligibility is placed on a continuum as to degree of intelligibility. The range is from unintelligible (not understood) to intelligible (fully understood). The listener assigns a rating (or score) related to the degree of intelligibility of a speech sample. The partitions most commonly used are 1-5, 1-7, and 1-9 points. According to Schiavetti, the purpose for the odd number scale is to provide a middle
value and two end points. Descriptive partitions can be assigned by the researchers, or numerical ratings can be chosen by the listener. In either case, the results tend to be the same (Schiavetti, 1992).

From the findings of their study, Schiavetti, Metz, and Sitler (1981) concluded equal-interval scaling should not be used as a measure of intelligibility because, as a variable, intelligibility is on a prothetic continuum (equal linear partitions cannot be assigned to the data) rather than a metathetic continuum (data are able to be partitioned into equal parts). Direct magnitude estimation falls in the metathetic continuum.

When using the method of direct magnitude estimation, the continuum is not constrained by either a maximum or minimum level. Direct magnitude estimation can be carried out in two ways. First, the listener hears a speech sample that has already been given a standard subjective value of intelligibility by the researcher. All subsequent samples are judged against this sample (Schiavetti, 1992). The other method of performing direct magnitude estimation is to have the listener assign any number to the first sample heard and rate the following samples as having a "perceived magnitude" of intelligibility based on the initial assignment (Schiavetti, 1992, p. 21). According to Schiavetti et al. (1981), although direct magnitude estimation has more construct validity than equal-interval
scaling in relation to intelligibility, they suggest further research needs to be completed on its criterion validity.

Both scaling procedures and write-down (orthographic transcription) methods yield the same type of information. According to Schiavetti et al. (1981), the reason scaling procedures often are chosen is they require less time and fewer listeners. According to Samar and Metz (1988), scaling procedures are considered by many to be "clinically tractable" (p. 307) by requiring less time and effort to administer. This was disputed in their study on intelligibility rating scale validity. While conducting their study that compared the rating-scale procedure to the write-down method, they were unable to find any "objective literature" (p. 315) in support of the reduced time factor of the scaling procedures in relation to the write-down method (Samar & Metz, 1988). In fact, the results of their study indicated that the two methods, (rating scale and write down) required approximately equal amounts of time to administer and score.

**Judgement Procedures.** Another type of intelligibility assessment often used by clinicians is the subjective measure of perceptual judgement. In this form of measurement, listeners assign a percentage of intelligibility based on their perception of how much of a speech sample they comprehend. Many times speech samples are obtained from speakers who are dysarthric, using word
lists or predetermined readings such as the "Rainbow" passage or "My Grandfather" (Beukelman & Yorkston, 1980; Darley, Aronson & Brown, 1969; Tikofsky, 1970). Yorkston and Beukelman (1980) suggested the purpose of this is to control some of the variables that can influence estimations of intelligibility, such as speech characteristic familiarity, type of speech sample taken, and contextual cues given to the listener. Using these standard passages, however, increases the risk of listener-sample familiarity, which limits the usefulness of estimation techniques for clinical assessment purposes (Beukelman & Yorkston, 1980).

Beukelman and Yorkston (1980) conducted a study that showed that speech-language pathologists consistently overestimate the level of intelligibility using the judgement procedure. Further investigation revealed that familiarity of the material (e.g., "My Grandfather") was a contributing factor to this consistent overestimation.

Scaling procedures and judgement procedures are two subjective measures of intelligibility frequently used by speech-language pathologists as noted by Gordon-Brannan (1993), Kent, Weismer, Kent, and Rosenbek (1989), Samar and Metz (1988), and Yorkston and Beukelman (1978). With measures such as these, validity always seems to be an issue and researchers continue to investigate their effectiveness.

There are also objective measures that quantify speech intelligibility. A discussion of some of these types of
procedures follows.

**Objective Measures of Intelligibility**

Objective measures yield numerical data for speech intelligibility. Methods vary as to form, but all attempt to move away from subjective judgements to a more concrete method of quantification.

**Word Identification Methods.** Two methods of word identification are open set and closed set. The open set word identification method consists of speech samples that are analyzed using the write-down or orthographic method of transcription. This procedure is considered the traditional method for determining percent of intelligibility and has been used by many prior investigators (Shriberg & Kwiatkowski, 1982; Weston & Shriberg, 1992; Yorkston & Beukelman, 1980). The format for this approach can be single words, sentences, or contextual speech samples. There are several advantages to this method. Results are typically reported as percentages, so they can be easily understood by both professionals and nonprofessionals. This type of measure (i.e., percentages) yield descriptive information that can be used to compare groups of related individuals, such as persons who are hearing impaired (Osberger, 1992) or dysarthric (Yorkston & Beukelman, 1978), and can be used to document gains as a result of treatment procedures (Gordon-Brannan, 1993; Schiavetti, 1992; Yorkston & Beukelman, 1978).
The orthographic method of transcription is the most widely used method of "quantifying intelligibility" (Yorkston & Beukelman, 1978, p. 499). This method also has a high degree of face validity (Kwiatkowski & Shriberg, 1992; Samar & Metz, 1988), but is considered to be more time-consuming in both administration and scoring, than some of the other more subjective measures such as scaling (Samar & Metz, 1988; Schiavetti, 1992; Subteiny, Van Hattum, & Meyers, 1972). Samar & Metz (1988) concluded from their study of both the orthographic method and the NTID (National Technical Institute for the Deaf) rating-scale procedure, that the orthographic procedure is not only more accurate, but more reliable for measuring intelligibility, especially in relation to the mid-range of speech intelligibility. According to Samar & Metz (1988), this is clinically significant as the majority of clients needing assessment will fall into this mid-range of intelligibility, as the extreme cases will be obvious. Although the orthographic method has been shown to be more accurate as a measure of intelligibility (Samar & Metz, 1988), it is important to note that this type of measure does not provide any information in regard to the etiology of the communication dysfunction (Osberger, 1992).

In this same category of objective measures are closed-set word identification tasks. These involve identification of words read from a word list. The P-SIM is an example of
this type of test.

P-SIM. The P-SIM (Wilcox et al., 1991) was designed to be used with preschoolers to obtain an objective measure of intelligibility. It is modeled after a standardized intelligibility test for adult dysarthric speakers developed by Yorkston and Beukelman (1982). In the P-SIM, twelve-word sets of homogeneous words are randomly selected for imitation. All words are one or two syllables in length. The child is asked to imitate 50 words, that are scored using the multiple choice format. Reliability of this measure was shown to be relatively stable. The P-SIM was shown to correlate highly with both the Goldman-Fristoe Test of Articulation (Goldman & Fristoe, 1986) and individual severity ratings assigned by speech-language pathologists and teachers (Wilcox et al., 1991).

Weiss Intelligibility Test. The Weiss Intelligibility Test is a standardized test that was developed to quantify intelligibility of both isolated words and contextual speech (Weiss, 1982). Independent scores from these two subtests are averaged to determine overall intelligibility. This published test is appropriate for use with both children and adolescents. Information is provided for determining factors contributing to unintelligibility. This test was standardized on 60 subjects aged 3 to 64 year. Although validity was established for this test using a sample of persons with varied disorders, further studies are
recommended to establish its validity among other various clinical populations. Administration and scoring is reported to be 10 to 15 minutes, dependent upon client age, cooperation, and degree of intelligibility. The Weiss Intelligibility Test is one example of the few standardized test instruments designed to specifically measure intelligibility.

Percentage of Consonants Correct (PCC). Designed by Shriberg and Kwiatkowski (1982), the PCC goes beyond intelligibility indexing and includes the constructs of disability and handicap. The PCC was designed for use with children whose articulation errors are predominantly deletions and substitutions. Severity of involvement is classified into four categories: mild, mild-moderate, moderate-severe, and severe. This method requires a 6-minute continuous speech sample. From this sample, consonant productions are scored as correct or incorrect, and the PCC value is calculated by dividing the number of correct consonants by the number of correct consonants plus the incorrect consonants. The result of this computation is multiplied by 100 to arrive at the PCC and the appropriate "perceived severity of involvement" (Shriberg & Kwiatkowski, 1982, p. 266). Further information can be obtained by analyzing the data score sheet that denotes specific error position in each consonant class. Shriberg & Kwiatkowski established construct validity, reliability, and clinical
utility for this measure. Distinct advantages to this method as stated by the authors are that it can be used repeatedly with the same child by one examiner for research or clinical purposes, and it does not require phonetic transcription but correct/incorrect judgements by the listener. The primary disadvantage of this method is that, if the sample is severely unintelligible, the listener may have a difficult time calculating a meaningful PCC. Consonant assessment can only occur if the listener understands the words intended by the child.

Children's Speech Intelligibility Test (CSIT). The CSIT word-recognition test, was designed by Kent et al. (1994) primarily for children with limited verbal abilities due to sensorimotor and cognitive deficits. Words are elicited using picture stimuli or imitation. Through careful selection of words, incorrect responses can be analyzed for specific error patterns (e.g., phonetic features). This information can then be compared to other speech-language developmental measures (Kent et al., 1994). The organization of this test is based on normal phonological development which permits the examiner to select test material appropriate to the child's ability. The "developmentally based construction" (Kent et al., 1994 p. 85) also allows for comparison of test results to normally developing same-aged peers. This test was designed to be used repeatedly throughout the course of treatment to
measure intelligibility progress. CSIT can also be used to analyze phonetic deficiencies that result in reduced intelligibility.

These are just a few of the objective measures used to quantify intelligibility. It is important to note that a complete assessment of intelligibility would include a baseline measure of percentage of words understood, and also an analysis of speech errors that contribute to the overall intelligibility of the client (Kent et al., 1990). Ultimately the method chosen will be determined by the procedure with which the examiner is familiar and comfortable and the purpose intended for the results (Yorkston & Beukelman, 1980).

Summary of Research Needs

A review of the literature indicates that, although there are many different ways to measure intelligibility, speech-language pathologists are still looking for an objective measure that is as valid as the orthographic method, but more time efficient. The Weiss check-slash method may be that measure. The moderate range of intelligibility was the focus of this study based on the fact that the majority of clients needing assessment will fall in this mid-range, as the extreme cases will be obvious (Samar & Metz, 1988).

In this study, the two methods that are most similar in the type of information gathered and in the form used to
achieve the goal will be compared. Many times clinicians want a valid objective measure for quantification of speech intelligibility to be used to determine baseline, chart progress, and compare groups. Both the orthographic measure and the check-slash method provide this type of information. The results of this research project can provide an alternative method of transcription for quantifying intelligibility of speech for individuals with disordered speech.
CHAPTER III

METHOD

This study compared the validity of a symbolic check-slash method of transcription designed by Weiss (1982) with the orthographic or manual transcription method to score continuous 100-word speech samples, elicited from children with moderate phonological deficits. The study also compared the time taken, recorded in minutes and seconds, to complete each procedure.

Subjects

The subjects for this study were 20 graduate students selected from the Speech and Hearing Sciences Graduate Program at Portland State University. Those selected met the following criteria:

1. Hold graduate student status in the Speech and Hearing Sciences Program for speech-language pathology.
2. Completion of one term of clinical experience working with children.
3. Normal bilateral hearing as measured by a pure tone audiometric hearing screening at 25 dB HL for the frequencies of 500, 1000, 2000, and 4000 Hz.

Subjects were recruited through graduate level classes in
the Speech and Hearing Sciences Program.

This study used data obtained from 4 of the original participants in the Gordon-Brannan study. These 4 subjects had transcribed the 100-word samples orthographically, following the same criteria as was set in this study. These 4 listeners did not keep track of transcription time, therefore, on the time data sheet (Appendix D) an asterisk (**) indicates no transcription times were available for these listeners. These four subjects were assigned to the orthographic transcription group and the remaining 16 subjects were randomly assigned to two groups. Group 1 transcribed the 12 speech samples using the orthographic method, and Group 2 used the check-slash transcription procedure.

Measuring Instruments

The hearing screening instrument used in this study was a portable AMBCO "Screen Ear", model 1122F, with Beltone Auraldome headphones. This is a wide range audiometer that utilizes the frequency range of 125-8000 Hz.

Procedures

Speech Samples

The speech samples used for transcription in this study were gathered by Gordon-Brannan (1993) in connection with her doctoral dissertation. The children were recruited from
local preschools and speech-language pathologists' case loads within the greater metropolitan area of Portland, Oregon. All of the children had no known neurological, motor, or physical impairments that could affect speech production. In addition, none exhibited any laryngeal or resonance deviancy at the time the samples were taken. Receptive comprehension was assessed using the Test of Auditory Comprehension of Language-Revised (TACL-R) (Carrow-Woolfolk, 1985), and only children who scored above the 10th percentile qualified for placement in the study. Hearing screenings were administered, and all but three children demonstrated normal bilateral hearing. Three displayed mild hearing losses, with pure tone averages of 35 dB or less bilaterally (Gordon-Brannan, 1993).

Of the original 48 connected speech samples, twelve 100-word continuous speech samples were used in this study. These samples had been elicited from 2 girls and 10 boys between the ages of 4:1 and 5:6 with a moderate degree of phonological deficiency.

The speech samples were gathered in an acoustically treated room especially designed for high-quality recordings. They were recorded using a Panasonic camcorder, VHS Recorder, AG-100, and a Sharp SX D200 digital audiotape recorder, in conjunction with an AKG, Model C451, capacitor flat response microphone. The speech samples were elicited using a book, The Relatives Came (Rylant & Gammell, 1985).
The 100-word speech sample was selected from the child telling a story using the book. Interjections, immediate repetitions of the investigator, and filler words (e.g., no, oh, naw, nope, yeah, and yup) were not included in the scored sample.

The severity rating for each child was determined using the 1-Minute Measure of Homonomy and Intelligibility designed by Hodson (1992). The phonological characteristics for the moderately deficient rating consisted of at least 12 phonemic substitutions, additions, and metathesis; and 1-9 omissions/glottal replacements (Gordon-Brannan, 1993). These speech samples were transcribed by four Portland State graduate students and were then analyzed by both Gordon-Brannan and the parents of the children for their accuracy. These verified transcriptions were used as the scoring key for this investigation.

Transcription Methods

The two methods of transcription used in this study were orthographic and check-slash. When using the orthographic style of manual transcription, the subjects wrote down word-for-word what they understood the child to say. Words that were not understood were denoted by using "X" as a place marker. Only complete words were counted as intelligible. Total transcription time was recorded for each sample to establish method efficiency (Appendix B).

The second method was a symbolic measure using a check
or a slash. This method uses a grid in which every box represents a single word expressed by the child (Appendix B). The listeners were instructed not to guess during this procedure. The listener marked each box with either a check (√) for each word understood (intelligible) or a slash (–) for words not understood (unintelligible). If there was any doubt as to the intelligibility of the utterance, it was considered not understood.

In order to provide context to the listeners, prior to listening to the audio tapes, they were given the book used to elicit the speech samples from the children.

Transcription Setting

Prior to listening to the speech samples, the subjects were given written instructions for transcribing the samples (Appendix C). Each listener had access to the eliciting materials (e.g., book) for reference while transcribing the audiotaped samples. The transcribing was completed in a room that had minimal distractions and individual work stations equipped with Sony model ER 9060 Educational recorders, in conjunction with Sony model 90 headphones. Each tape player was equipped with variable speed control that allowed the subjects to transcribe the samples at their own pace.

Transcription Procedure

Participants were required to self-monitor their
listening experience. Each subject was allowed to listen to the audiotaped speech samples a maximum of three times. Utterances could be listened to one at a time. Once the subject's transcription was completed, the listeners had the option of listening to the whole sample one time to verify their transcriptions. All samples were numbered 1-12 for identification purposes. Each listener used a stopwatch to record the amount of time used to complete each transcription. Completed transcriptions were collected by the researcher and scored.

Reliability

Interrater reliability was established using the Pearson product-moment correlation coefficient (Pearson r). This score was determined by averaging the results of each of the Pearson r values for the 45 possible pairs of listeners. In addition, a reliability analysis was computed using a generalization of Cohen's kappa statistic that measures agreement between two observers. Berry and Mielke (1988) designed this statistic to extend Cohen's kappa, to both interval level data and multiple raters. Pairs of listeners were compared for agreement across samples. The multirater version of the Cohen kappa statistic was chosen to demonstrate how closely the raters were in agreement with one another on the percentage of intelligibility score assigned to each sample. Intra-rater reliability was not established, based on the high probability of listener
familiarity, due to multiple opportunities for each subject to listen to each sample.

Scoring and Data Analysis

Orthographic transcriptions for each 100-word sample for each listener were scored using the transcript scoring key. To be considered correct, transcribed words had to be a direct match to the key or differing only in morphological form. Each sample was assigned a percentage of intelligibility, determined by the number of correctly matched words with the 100-word scoring key. The check-slash transcriptions for each sample and listener were assigned a percentage of intelligibility by dividing the number of intelligible words by the total number of unintelligible and intelligible words recorded using the symbol system. Descriptive statistics for intelligibility data for each transcription method were computed for each sample for comparison, indicating the median and interquartile range. These two measures were chosen because of the nonsymmetrical nature of the data. Due to the skewed nature of the data and the small sample size, each sample and method were compared for statistical significance using the nonparametric, two-tailed Mann-Whitney U Test for independent samples. The p-value was set at <.10 to control for type II errors.

Descriptive statistics for the amount of transcription
time necessary for each method, were compiled indicating the
group median and interquartile range for each sample.
Transcription times for each sample were analyzed for
significant differences using a one-tailed nonparametric
Mann-Whitney U Test for independent samples. This test was
chosen due to the non-normal distribution of the population
and the relatively small sample size.
CHAPTER IV

RESULTS AND DISCUSSION

Results

The purpose of this study was to determine the validity and efficiency of the check-slash method of transcription for determining percentage of intelligibility, as compared to the orthographic transcription method. The first question addressed in this research project was to determine if the check-slash method of transcription is valid when used to assess individuals with a moderate phonological deficit. This was accomplished by comparing intelligibility scores (stated in percentages) obtained for 12 connected speech samples transcribed using the orthographic and check-slash methods of transcription. (Raw data appear in Appendix D.)

Reliability Analysis

Interjudge reliability for the listeners was assessed using the Pearson $r$ and a multirater reliability analysis. The mean Pearson $r$ for the check-slash method was .80, and .79 for the orthographic method. Tables 1 and 2 provide correlation coefficient data for interrater reliability for the check-slash and orthographic transcription methods,
Table 1

**Correlation Matrix for Comparison of Listeners by Pairs for the Check-Slash Method of Transcription across the 12 Samples**

<table>
<thead>
<tr>
<th></th>
<th>L1</th>
<th>L2</th>
<th>L3</th>
<th>L4</th>
<th>L5</th>
<th>L6</th>
<th>L7</th>
<th>L8</th>
<th>L9</th>
<th>L10</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2</td>
<td>.83*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L3</td>
<td>.92*</td>
<td>.86*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L4</td>
<td>.90*</td>
<td>.83*</td>
<td>.91*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L5</td>
<td>.80*</td>
<td>.92*</td>
<td>.90*</td>
<td>.85*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L6</td>
<td>.66*</td>
<td>.87*</td>
<td>.85*</td>
<td>.77*</td>
<td>.86*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L7</td>
<td>.46*</td>
<td>.78*</td>
<td>.68*</td>
<td>.64*</td>
<td>.78*</td>
<td>.92*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L8</td>
<td>.58*</td>
<td>.60*</td>
<td>.69*</td>
<td>.72*</td>
<td>.73*</td>
<td>.73*</td>
<td>.61*</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L9</td>
<td>.78*</td>
<td>.91*</td>
<td>.89*</td>
<td>.86*</td>
<td>.91*</td>
<td>.92*</td>
<td>.81*</td>
<td>.82*</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>L10</td>
<td>.94*</td>
<td>.92*</td>
<td>.91*</td>
<td>.93*</td>
<td>.74*</td>
<td>.61*</td>
<td>.59*</td>
<td>.84*</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

**Check-Slash Mean Correlation Between Judges**

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.80</td>
<td>.46</td>
<td>.94</td>
<td>.11</td>
</tr>
</tbody>
</table>

Note: L = Listener; Marked correlations (*) are significant at p < .05.
Table 2

Correlation Matrix for Comparison of Listeners by Pairs for the Orthographic Method of Transcription across the 12 Samples

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2</td>
<td>.70*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L3</td>
<td>.65*</td>
<td>.75*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L4</td>
<td>.77*</td>
<td>.87*</td>
<td>.89*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L5</td>
<td>.79*</td>
<td>.92*</td>
<td>.78*</td>
<td>.85*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L6</td>
<td>.72*</td>
<td>.67*</td>
<td>.67*</td>
<td>.77*</td>
<td>.72*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L7</td>
<td>.75*</td>
<td>.89*</td>
<td>.70*</td>
<td>.78*</td>
<td>.97*</td>
<td>.75*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L8</td>
<td>.74*</td>
<td>.89*</td>
<td>.88*</td>
<td>.85*</td>
<td>.92*</td>
<td>.79*</td>
<td>.91*</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L9</td>
<td>.69*</td>
<td>.89*</td>
<td>.63*</td>
<td>.73*</td>
<td>.94*</td>
<td>.67*</td>
<td>.96*</td>
<td>.88*</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>L10</td>
<td>.66*</td>
<td>.75*</td>
<td>.77*</td>
<td>.72*</td>
<td>.82*</td>
<td>.77*</td>
<td>.87*</td>
<td>.88*</td>
<td>.77*</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Orthographic Mean Correlation Between Judges

<table>
<thead>
<tr>
<th>Mean</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>.77</td>
<td>.62</td>
<td>.97</td>
<td>.09</td>
</tr>
</tbody>
</table>

Note: L = Listener; Marked correlations (*) are significant at p < .05.
respectively. These values indicate that, within each method, there was a strong linear association between the listeners, across the samples. However, this does not necessarily mean that the listeners were in agreement in the percentages assigned, but that the proportional differences between samples was consistent. For this reason, the Pearson \( r \) was not enough to substantiate interrater agreement in actual percentages assigned. Therefore, a second method of establishing reliability was used. The results of the multirater kappa measure of agreement for the orthographic method was .40 and the check-slash was .23. These results indicate that the interrater agreement for both methods was low.

**Intelligibility Analysis**

To determine if there was a significant difference in the percentages of intelligibility obtained when using the check-slash or the orthographic transcription methods, intelligibility percentage scores were determined for each of the 12 samples from the transcriptions (orthographic or check-slash) of all 20 listeners and then analyzed (Appendix D). A Mann-Whitney U Test was performed on each of the 12 samples to compare the two methods for significant differences.

Review of the data revealed that the difference between median values of the two methods for all samples was relatively small at 8 or less percentage points, with the
exception of sample #12 at 15 percentage points. Comparisons of the medians and interquartile ranges for both methods are presented in Table 3 along with the z and p values from the Mann-Whitney U Tests.

The results from the 12 Mann-Whitney U Tests indicated no significant difference between the check-slash and orthographic transcription methods when determining percentage of intelligibility for individuals with a moderate phonological impairment. For this study the p-level was considered significant at p < .10 in order to control for type II error, that is acceptance of a false null hypothesis. Type II errors are normally controlled for by increased sample sizes (larger than 30 subjects). Due to the relatively small sample size of this study, it was important to control for this type of error. The power of the Mann-Whitney U Test was equal to .68 which means that if there was a typical size effect (a moderate size difference between the two transcription methods), there was a 68% chance that it would have been detected by the Mann-Whitney U Tests performed.

**Time Analysis**

The second question addressed in this study was to determine if the check-slash method of transcription was a more time-efficient procedure when compared to the orthographic method. A one-tailed Mann-Whitney U test was used to analyze each of the 12 samples for significant
Table 3

Descriptive Statistics for Percentage of Intelligibility. Comparison of the Check-Slash and Orthographic Methods of Transcription N=12

<table>
<thead>
<tr>
<th></th>
<th>Check Median</th>
<th>Check I-Q Range</th>
<th>Ortho Median</th>
<th>Ortho I-Q Range</th>
<th>Diff. Medians</th>
<th>M-W z</th>
<th>M-W p</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>72</td>
<td>59 - 78</td>
<td>77</td>
<td>72 - 83</td>
<td>5</td>
<td>1.55</td>
<td>.12</td>
</tr>
<tr>
<td>#2</td>
<td>73</td>
<td>61 - 83</td>
<td>75</td>
<td>68 - 78</td>
<td>2</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>#3</td>
<td>70</td>
<td>56 - 75</td>
<td>63</td>
<td>49 - 68</td>
<td>7</td>
<td>1.40</td>
<td>.16</td>
</tr>
<tr>
<td>#4</td>
<td>81</td>
<td>71 - 86</td>
<td>82</td>
<td>79 - 85</td>
<td>1</td>
<td>.64</td>
<td>.52</td>
</tr>
<tr>
<td>#5</td>
<td>78</td>
<td>67 - 85</td>
<td>73</td>
<td>70 - 74</td>
<td>5</td>
<td>.94</td>
<td>.34</td>
</tr>
<tr>
<td>#6</td>
<td>59</td>
<td>45 - 65</td>
<td>64</td>
<td>50 - 66</td>
<td>5</td>
<td>.72</td>
<td>.47</td>
</tr>
<tr>
<td>#7</td>
<td>80</td>
<td>60 - 84</td>
<td>76</td>
<td>68 - 81</td>
<td>4</td>
<td>.42</td>
<td>.68</td>
</tr>
<tr>
<td>#8</td>
<td>81</td>
<td>73 - 83</td>
<td>83</td>
<td>74 - 84</td>
<td>2</td>
<td>.41</td>
<td>.68</td>
</tr>
<tr>
<td>#9</td>
<td>41</td>
<td>28 - 49</td>
<td>49</td>
<td>43 - 53</td>
<td>8</td>
<td>1.36</td>
<td>.17</td>
</tr>
<tr>
<td>#10</td>
<td>86</td>
<td>70 - 87</td>
<td>84</td>
<td>79 - 88</td>
<td>2</td>
<td>.49</td>
<td>.62</td>
</tr>
<tr>
<td>#11</td>
<td>84</td>
<td>74 - 87</td>
<td>79</td>
<td>73 - 82</td>
<td>5</td>
<td>1.51</td>
<td>.13</td>
</tr>
<tr>
<td>#12</td>
<td>53</td>
<td>59 - 72</td>
<td>68</td>
<td>64 - 72</td>
<td>15</td>
<td>1.24</td>
<td>.21</td>
</tr>
</tbody>
</table>

Note: S = Sample; Check = Check-Slash Transcription Method; Ortho = Orthographic Transcription Method; Diff. = Difference; M-W = Mann-Whitney; I-Q = Interquartile; * = significant at p < .10. Differences in the amount of time necessary to transcribe
the samples using the two methods.

The results of the Mann-Whitney U Tests revealed that for 11 of the 12 samples, there was a significant difference in the amount of time needed to transcribe the samples with $p < .05$. In all instances, the median time to transcribe the samples was less for the check-slash method. The increased efficiency varied from 38% to 54%. Median transcription times for the check-slash method ranged from 3.39 to 8.65 minutes. The orthographic median time ranged from 7.22 to 13.10 minutes. Median scores and interquartile ranges for both procedures are represented in Table 4, along with the Mann-Whitney U Test $z$ and $p$ values.

In summary, the results from this study indicate that, although both methods had low overall reliability for actual listener agreement (.40 for the orthographic and check-slash at .23), the Pearson $r$ correlation coefficients for interrater reliability were strong at .80 & .79,. There was no significant difference found between the two methods in determining percent of intelligibility. Results of the study established the overall efficiency of the check-slash method when compared to the orthographic transcription method.

Discussion

This investigation sought to establish validity for the check-slash method of transcription designed by Weiss (1982)
Table 4

**Descriptive Statistics: Transcription Time Required to Complete each Sample using Both Methods**

<table>
<thead>
<tr>
<th>S</th>
<th>Check Median Time</th>
<th>Check I-Q Range</th>
<th>Ortho Median Time</th>
<th>Ortho I-Q Range</th>
<th>M-W z</th>
<th>M-W p</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>8.65</td>
<td>4.42 - 9.49</td>
<td>13.10</td>
<td>6.95 - 13.44</td>
<td>1.95</td>
<td>.025*</td>
</tr>
<tr>
<td>#2</td>
<td>6.61</td>
<td>4.33 - 8.65</td>
<td>10.03</td>
<td>7.02 - 12.05</td>
<td>1.74</td>
<td>.041*</td>
</tr>
<tr>
<td>#3</td>
<td>5.15</td>
<td>4.85 - 8.70</td>
<td>10.77</td>
<td>6.37 - 12.28</td>
<td>2.17</td>
<td>.015*</td>
</tr>
<tr>
<td>#4</td>
<td>4.63</td>
<td>3.67 - 6.48</td>
<td>8.80</td>
<td>5.50 - 11.73</td>
<td>2.49</td>
<td>.006*</td>
</tr>
<tr>
<td>#5</td>
<td>4.71</td>
<td>3.32 - 5.88</td>
<td>8.69</td>
<td>5.37 - 8.77</td>
<td>2.60</td>
<td>.005*</td>
</tr>
<tr>
<td>#6</td>
<td>6.43</td>
<td>3.98 - 8.39</td>
<td>10.72</td>
<td>5.65 - 8.17</td>
<td>2.06</td>
<td>.019*</td>
</tr>
<tr>
<td>#7</td>
<td>4.74</td>
<td>3.42 - 6.90</td>
<td>12.54</td>
<td>5.50 - 9.63</td>
<td>2.28</td>
<td>.011*</td>
</tr>
<tr>
<td>#8</td>
<td>4.80</td>
<td>3.25 - 6.84</td>
<td>7.80</td>
<td>3.08 - 9.13</td>
<td>1.36</td>
<td>.086</td>
</tr>
<tr>
<td>#9</td>
<td>6.33</td>
<td>3.22 - 6.89</td>
<td>10.28</td>
<td>5.68 - 11.33</td>
<td>1.84</td>
<td>.033*</td>
</tr>
<tr>
<td>#10</td>
<td>5.84</td>
<td>5.13 - 6.60</td>
<td>8.63</td>
<td>6.95 - 9.17</td>
<td>2.49</td>
<td>.006*</td>
</tr>
<tr>
<td>#11</td>
<td>3.39</td>
<td>2.75 - 5.85</td>
<td>7.22</td>
<td>5.57 - 8.06</td>
<td>2.28</td>
<td>.011*</td>
</tr>
<tr>
<td>#12</td>
<td>5.26</td>
<td>3.10 - 6.29</td>
<td>8.43</td>
<td>5.25 - 9.13</td>
<td>2.17</td>
<td>.015*</td>
</tr>
</tbody>
</table>

Note: S = sample; Check = Check-Slash Transcription Method; Ortho = Orthographic Transcription Method; M - W = Mann-Whitney U Test; I-Q = Interquartile; Significant p values (p < .05) marked with (*).
when used to calculate a percentage of intelligibility to individuals with moderate phonological speech impairments. The orthographic method was used for comparison as it is considered by many speech-language pathologists to have high face validity (Beukelman & Yorkston, 1979; Kwiatkowski & Shriberg, 1992; Samar & Metz, 1988).

**Intelligibility**

This study failed to establish validity for the check-slash method of transcription when used to assess the speakers with moderate phonological deficiencies. In this study, the result of the Pearson $r$ for the check-slash method was .80, and .79 for the orthographic method. This indicates a strong linear association between the listeners, across the samples. However, this does not necessarily mean that the listeners were in agreement in the percentages assigned, but that the proportional differences between samples was consistent. For this reason a second method of reliability was computed.

A more appropriate means of assessing reliability among multiple raters has been developed by Berry and Mielke (1988). This multirater kappa is a generalization of Cohen's statistic kappa. Cohen's kappa is a "reliability index for measuring chance-corrected agreement between two observers employing nominal scales" (p. 921) (Berry & Mielke, 1988). Berry and Mielke extended Cohen's kappa to both interval level data and multiple raters. Like Cohen's
kappa, the multirater kappa is chance-corrected and utilizes "Euclidean distances" (Berry & Mielke, 1988). This multirater kappa is used to assess "the degree to which judges agree on their scoring, above and beyond what is expected by chance" (Berry & Mielke, 1988, p.927). Analysis of data using the multirater kappa result in a single "coefficient of agreement" that ranges from .00 being equal to chance and 1.00 equal to total agreement. A multirater kappa analysis for each method was computed. The measure of reliability for both the check-slash transcription method and the orthographic method was low (.23 for the check-slash method and .40 for the orthographic method). This result is due in part to the high amount of variability of percentage of intelligibility calculated for each sample across listeners (Appendix D). An extreme example of this can be seen in sample #3. Within this sample, the calculated percent of intelligibility for the orthographic group ranged from 46% to 79% (33 percentage points), and the check-slash range was from 38% to 82% (44 percentage points). In contrast, sample #10 had much smaller ranges with the orthographic range at 76% to 89% (13 percentage points), and the check-slash range at 68% to 91% (23 percentage points). Because the samples were all similar in the respect that they were all classified as moderately phonologically impaired in the Gordon-Brannan study (1993), it was expected that the ranges for each sample would be small. These large
ranges demonstrate the variability of the samples, individuality of the listeners, and the difficulty of the task.

The homogeneous nature of the sample studied also contributes to the low overall method reliability. As stated by Samar & Metz (1988), it is this moderate range of intelligibility that is difficult to assess, as the extreme cases are obvious. This can best be understood by placing intelligibility on a continuum from normal (100% intelligible) to profound (100% unintelligible). If this study had included both the mildly and profoundly impaired, it is probable that the reliability for those samples would have been much higher which would have pulled the overall reliability for both methods up considerably.

The task of transcribing language samples is subjective in nature. The result is dependent on the listener's ability to identify individual phonemes and assign meaning. Each listener comes to the task with varying amounts of experience and skill. The results of this study indicate that the listeners who participated had varying levels of perceptual skill. This would partially account for the large differences in intelligibility ratings.

It is important to note that overall the listeners were consistent in their ability to rate the samples, as is indicated by the strong Pearson r scores. The data clearly illustrate that some listeners consistently understood a
higher percentage of the words in each of the 12 samples, while others consistently understood a lower percentage of words. Data collected from listeners #9 and #12 are an illustration of this (Appendix D). Percentage of intelligibility for the 12 samples judged by listener #9 were consistently higher than median, and intelligibility scores for listener #12 consistently were lower than the median. The difference between intelligibility scores for these two listeners was consistently 20 to 30 percentage points across all 12 connected speech samples. This pattern of consistency led to the high Pearson r correlation coefficient.

Additionally, listeners who tended to score samples as highly unintelligible, generally scored many samples lower than the median. This can be seen in orthographic listener #6 and check-slash listeners #12 & 19 (Raw Data Chart, Appendix D). This trend was also true for those listeners that scored the samples higher than the median, (listeners #9, #16 and #18).

Although the nonparametric analysis indicated that there is no significant difference between these two methods when used to assign a percentage of intelligibility to individuals with moderate phonological impairments, the raw data chart (Appendix D) shows that the overall variability is consistently larger within the check-slash method of transcription. This is more accurately reflected in the
interquartile ranges for each sample and method (Table 3). The interquartile range eliminates the "outliers" (those scores that are at either extremes) and give a truer representation of the variation between methods. It is interesting to note that the medians for both methods are within ±8 percentage points for all samples excluding sample 12, and 9 samples out of 12 were within ±5 percentage points.

The results of this study should not be interpreted to mean that these methods are equal in determining an accurate percentage of intelligibility, but rather that there is no significant difference between their clinical application in calculating a percentage of intelligibility. The accuracy of the percentage of intelligibility obtained would be dependent on the perceptual skill of the listener. Further research needs to be conducted with a larger sample to determine if these methods are equivalent in their clinical application.

**Time**

According to the results of this study, as expected there was a significant difference in the amount of time necessary to transcribe the language samples using the two transcription methods. The results indicate that check-slash method of transcription was far more efficient.

The issue of time is often addressed in the literature when discussion is focused on the inefficiency of the
orthographic transcription method. Many authors have stated that language samples are often assigned a subjective percentage of intelligibility rating, rather than being transcribed, primarily due to the amount of time orthographic transcription requires (Morrison & Shriberg, 1992; Samar & Metz, 1988; Schiavetti, 1992; Subtelny et al., 1972). This study supports those statements by showing that the orthographic method took from 38% to 54% more time to complete (Table 4). The overall range for the median times needed to transcribe the 12 samples using the check-slash methods was 3.39 to 8.65 minutes (Table 4). The orthographic transcription medians ranged from 7.22 to 13.10 minutes. The 38% time difference can be found in sample #12 with the check-slash median time at 5.26 minutes and the orthographic median at 8.43 minutes. The largest difference was in sample #11 with the check-slash median time recorded as 3.39 minutes and the orthographic median at 7.22 minutes. Sample #8 was the only sample that was nonsignificant (Table 4). An accurate reflection of the differences in the time needed to transcribe the 12 samples is represented in the interquartile ranges and medians (Table 4) as the outliers are eliminated.

This result is not surprising as one would assume it would take more time to write word-for-word what has been said than it would to record a symbol for each word uttered. The transcription times recorded in this study indicate
large amounts of variability due to listener individuality. For example, the orthographic range for sample #1 was 6.95 to 23.77 minutes, and the check-slash range extended from 4.30 to 12.42 minutes (Appendix D). This discrepancy in the amount of time necessary to transcribe the samples could account for some of the variability within the intelligibility scores. If listeners hastily transcribed a sample, then the percentage of intelligibility could have been lower than if they had taken more time. A closer look at the data reveals that the listeners who consistently arrived at lower levels of intelligibility (Listeners 6, 12, and 19) spent less time than the median to transcribe the samples, while check-slash Listener 16 and orthographic Listener 9 took longer than the median for most samples and arrived at a percent of intelligibility higher than median.

The conclusion that can be drawn from this study is that the check-slash method of transcription has the potential for greater efficiency, when used to calculate a percentage of intelligibility. The amount of time will vary dependent on two factors: (a) the severity of the client, with the mild and severe-profound being the quickest to transcribe (based on the likelihood that the amount of intelligibility is readily identifiable) and the moderately impaired requiring more time (Samar & Metz, 1988), and (b) the experience of the clinician using the method. It can be assumed that those who use a method frequently will develop
proficiency.
CHAPTER V

SUMMARY AND IMPLICATIONS

Summary

The field of speech-language pathology is currently in a state of change. Reimbursement for services provided to clients requires documentation in all aspects of diagnosis and treatment progress. Numerical data provide stronger support for clinical diagnosis and progress over other subjective measures. Efficient use of billable time for services provided continues to be an issue. Therefore, it is important that speech-language pathologists find clinical methods of documentation that are both valid and time-efficient.

The purpose of this study was to provide validity for a check-slash method of transcription, used for determining percentage of intelligibility of speakers who are moderately phonologically impaired. The orthographic method of transcription was chosen for comparison because of its established high face validity. A second purpose of this research project was to establish the efficiency of the check-slash method by comparing the amount of time necessary to transcribe language samples using both the check-slash and orthographic transcription methods.
The data were collected from 20 graduate students randomly assigned to two transcription groups (check-slash and orthographic). Spontaneous 100-word speech samples gathered from 2 girls and 10 boys with moderate phonological impairments, aged 4:1 to 5:6, were used for transcription. One group used the check-slash method of transcription and the other used the orthographic method. The listeners in each group transcribed each sample and kept track of the amount of time taken to complete each sample. This researcher scored each sample and assigned a percentage of intelligibility based on transcription results.

Validity for the check-slash method of transcription was not established by the results of this study. Although results indicate no significant difference between the two methods when used to assign a percentage of intelligibility, the multirater kappa for interrater reliability of both methods was weak (.23 for the check-slash transcription method and .40 for the orthographic method). This low reliability was due primarily to the large amount of variability within samples, which is a result of differences in individual listener perceptions. This nonsignificant result does not mean that the reliability of the two methods is equal, as the data reflect that the variation was greater in the check-slash method of transcription. The difference in the interrater mean reliability is significant as it is reflective of the large amount of variability demonstrated
in the check-slash group which resulted in the lower mean reliability score.

The efficiency of the check-slash transcription method in comparison to the orthographic method was established by this study. Overall the check-slash transcription method was 38% to 54% more time-efficient than the orthographic method. The data reflect that listeners who took the least amount of time to transcribe the samples often recorded intelligibility scores that were lower than the median. In contrast, those listeners, in both groups, who took the longest to transcribe the samples tended to record intelligibility scores higher than the medians.

Implications

Clinical

This study has brought to question the reliability of both the check-slash and orthographic methods of transcription. The results of this study demonstrate how each clinician comes to the task with varying levels of skill in this area. Some clinicians appear to be intuitively better at the task.

This researcher believes that the check-slash method of transcription could be a clinically useful tool. This study showed that listeners are fairly consistent in the way they score language samples for percentage of intelligibility. In this study listeners tended to assign either high or low
intelligibility scores to the samples. This consistency would be the basis for utilizing this method. When used in a consistent manner, percentage of intelligibility obtained with the check-slash method would probably be very similar to the percentage obtained through orthographic transcription from the same listener, as listener perception would remain constant.

The clinical implication for time is obvious and needs minimal discussion. This method has been shown to be significantly more efficient than the orthographic method. The efficiency of this method would encourage clinicians to obtain numerical support for the percentage of intelligibility they assign to individuals with speech impairments when diagnosing and making progress statements. Additionally, this method has the potential to be used "on line", unlike the orthographic method which is much more cumbersome.

Research

Further research involving the check-slash transcription method would continue to strive to establish validity for the method. This could be attempted by increasing the sample size to greater than 30 listeners. It would also be interesting to have the same listener transcribe two different language samples, taken from the same speaker, using both methods. This would demonstrate the effects of listener perception while showing the true
correlation between the check-slash and orthographic transcription methods.

The variability in percentage of intelligibility was very high in this study. Possibly the ranges of percent intelligibility would have been narrowed if skilled practicing speech-language pathologists were the listeners in this study rather than graduate students.

As the check-slash method has been proven to be an efficient measure, future research should be pursued to establish necessary validity to make this method a useful clinical tool.
References


Hodson, B. W., & Paden, E. P. (1981). Phonological processes which characterize unintelligible and intelligible
speech in early childhood. Journal of Speech and Hearing Disorders, 46, 369-373.


Journal of Speech and Hearing Research, 13, 59-64.


APPENDIX A

FACTORS THAT AFFECT INTELLIGIBILITY
Factors that Affect Intelligibility

The following are factors that affect speaker intelligibility. This list was compiled by Weiss (1982) as an optional section of the Weiss Intelligibility Test. Each category is rated using a scale comprised of: normal, mildly abnormal, moderately abnormal, severely abnormal and profoundly abnormal.

- Adventitious Sounds
- Articulation
- Communicative Disfluency
- Disfluency Redundancy
- Inflection
- Juncture
- Mean Length of Utterance
- Morphology
- Morphophonemics
- Pauses
- Physical Posture
- Pitch

Pragmatics
Pronunciation
Rate
Resonation
Rhythm
Semantics
Stress
Syntax
Voice Quality

APPENDIX B

TRANSCRIPTION FORMS
CHECK-SLASH TRANSCRIPTION FORM

TRANSCRIBER: ____________________________________________________________

SAMPLE NUMBER: _________________________________________________________

DIRECTIONS:
Please write your name and the sample number you are working on in the space provided above. Code the sample on the grid provided below. For each word you definitely understand, place a check (✓) in a square. For each word you definitely did not understand place a slash (−) in the square. Please do not guess. Code only the child's utterances. Please, do not code any interjections, immediate repetitions of the investigator or filler words (e.g., no, oh, naw, nope, yeah, and yup).

Listen to the sample stopping as needed to write. You may listen to each utterance a maximum of three times. When you have finished you may listen to the tape one last time to verify your transcription.

START YOUR STOPWATCH TO RECORD YOUR TOTAL TRANSCRIPTION TIME BEFORE YOU BEGIN AND RECORD YOUR STOP TIME BELOW IN THE SPACE PROVIDED.

TOTAL TRANSCRIPTION TIME FOR SAMPLE: _____ Min. _____ Sec
ORTHOGRAHIC TRANSCRIPTION FORM

TRANSCRIBER: ____________________________

SAMPLE NUMBER: __________________________

Directions:
Please write your name and the sample number you are working on in the space provided above. Use the space provided below to transcribe the 100-word speech sample. You may not need all the spaces provided. Write down each word you understand the child to say. Code only the child's utterances. Please, do not code any interjections, immediate repetitions of the investigator or filler words (e.g., no, oh, naw, nope, yeah, and yup). For each word not understood use an "X" as a place marker.

Listen to the sample stopping as need to write. You may listen to each utterance a maximum of three times. When you have finished you may listen to the tape one last time to verify your transcription. Record total transcription time in the space provided below.

START YOUR STOPWATCH TO RECORD TRANSCRIPTION TIME BEFORE YOU BEGIN.

ORTHOGRAHIC TRANSCRIPTION SHEET

1. ____________________________
2. ____________________________
3. ____________________________
4. ____________________________
5. ____________________________
6. ____________________________
7. ____________________________
8. ____________________________
9. ____________________________
10. ____________________________
11. ____________________________
12. ____________________________
13. ____________________________
14. ____________________________
15. ____________________________
16. ____________________________
17. ____________________________
18. ____________________________
19. ____________________________
20. ____________________________

TOTAL TRANSCRIPTION TIME FOR SAMPLE: _____ Min. _____ Sec
APPENDIX C

TRANSCRIPTION INSTRUCTIONS
Transcription Instructions

You will be listening to twelve 100-word continuous speech samples taken from children 4:1 - 5:6 years of age with a moderate phonological deficiency. Please familiarize yourself with the materials provided for context prior to listening to the tape.

You are to transcribe these samples using the orthographic method of transcription. Procedures for this method follow.

Instruction for orthographic transcription.

Before beginning, please write your name and the sample number in the spaces provided, then start your stopwatch. Work on one sample at a time. Listen to the sample, stopping as frequently as needed to write. Write down word-for-word what you hear the child say. Any words that you cannot understand mark with "X" as a place marker. Do not transcribe the following: clinician's responses, interjections, immediate repetitions of the clinician or any filler words (e.g., no, oh, naw, nope, yeah, and yup). You may listen to each utterance a maximum of three times for this procedure. When you are finished, you may listen to the entire sample one last time to verify your transcription. When finished stop your stopwatch.

Please note the total transcription time for each sample by recording the stop-time in minutes and seconds at the bottom of your transcription sheet in the space
provided.

If you have any questions please contact me. Please return all materials and completed transcriptions to this researcher.

Thank you for your participation!

Vicky Bacon
**Transcription Instructions**

You will be listening to twelve 100-word continuous speech samples taken from children 4:1 - 5:6 years of age with a moderate phonological deficiency. Please familiarize yourself with the materials provided for context prior to listening to the tape.

You are to transcribe these samples using the check-slash method of transcription. Procedures for this method follow.

**Instruction for check-slash method of transcription.**

Before beginning, please write your name and the sample number in the spaces provided, then start your stopwatch. Work on one sample at a time, stopping as frequently as needed to mark the transcription grid provided. Code only the child's utterances. **Do not code the clinician's responses,** interjections, immediate repetitions of the clinician or any filler words (e.g., no, oh, naw, nope, yeah, and yup). On the transcript grid, each box represents a word. Mark a check (√) for each word you understand, and place a slash (-) in each space for any word you do not understand. If there is any question as to if you understand a word, mark it as unintelligible by using a slash (-) . **PLEASE DO NOT GUESS.** You may listen to each utterance a maximum of three times for this procedure. When you have finished you may listen to the tape one last time to verify your transcription. When you are finished with
each transcription stop your stopwatch.

Please note the total transcription time for each sample by recording the stop-time in minutes and seconds at the bottom of your transcription sheet in the space provided.

If you have any questions please contact me. Please return all materials and completed transcriptions to this researcher.

Thank you for your participation!

Vicky Bacon
APPENDIX D

RAW DATA
### Percentage of Intelligibility per Sample per Judge

<table>
<thead>
<tr>
<th>Judge</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S5</th>
<th>S6</th>
<th>S7</th>
<th>S8</th>
<th>S9</th>
<th>S10</th>
<th>S11</th>
<th>S12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ortho</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>88</td>
<td>68</td>
<td>68</td>
<td>89</td>
<td>61</td>
<td>49</td>
<td>70</td>
<td>74</td>
<td>53</td>
<td>85</td>
<td>82</td>
<td>67</td>
</tr>
<tr>
<td>2</td>
<td>77</td>
<td>77</td>
<td>67</td>
<td>85</td>
<td>72</td>
<td>69</td>
<td>82</td>
<td>84</td>
<td>51</td>
<td>88</td>
<td>74</td>
<td>63</td>
</tr>
<tr>
<td>3</td>
<td>84</td>
<td>71</td>
<td>46</td>
<td>81</td>
<td>74</td>
<td>62</td>
<td>83</td>
<td>84</td>
<td>53</td>
<td>83</td>
<td>76</td>
<td>76</td>
</tr>
<tr>
<td>4</td>
<td>89</td>
<td>67</td>
<td>56</td>
<td>87</td>
<td>75</td>
<td>64</td>
<td>80</td>
<td>78</td>
<td>44</td>
<td>87</td>
<td>70</td>
<td>69</td>
</tr>
<tr>
<td>5</td>
<td>77</td>
<td>73</td>
<td>65</td>
<td>79</td>
<td>70</td>
<td>60</td>
<td>75</td>
<td>83</td>
<td>43</td>
<td>79</td>
<td>79</td>
<td>64</td>
</tr>
<tr>
<td>6</td>
<td>62</td>
<td>56</td>
<td>49</td>
<td>78</td>
<td>70</td>
<td>50</td>
<td>58</td>
<td>53</td>
<td>37</td>
<td>76</td>
<td>73</td>
<td>61</td>
</tr>
<tr>
<td>7</td>
<td>74</td>
<td>78</td>
<td>70</td>
<td>82</td>
<td>72</td>
<td>65</td>
<td>74</td>
<td>83</td>
<td>46</td>
<td>83</td>
<td>82</td>
<td>70</td>
</tr>
<tr>
<td>8</td>
<td>72</td>
<td>78</td>
<td>52</td>
<td>83</td>
<td>73</td>
<td>54</td>
<td>77</td>
<td>81</td>
<td>43</td>
<td>83</td>
<td>78</td>
<td>66</td>
</tr>
<tr>
<td>9</td>
<td>82</td>
<td>90</td>
<td>79</td>
<td>85</td>
<td>80</td>
<td>72</td>
<td>83</td>
<td>86</td>
<td>56</td>
<td>89</td>
<td>85</td>
<td>74</td>
</tr>
<tr>
<td>10</td>
<td>73</td>
<td>78</td>
<td>61</td>
<td>81</td>
<td>75</td>
<td>67</td>
<td>67</td>
<td>85</td>
<td>56</td>
<td>88</td>
<td>82</td>
<td>76</td>
</tr>
<tr>
<td>Check</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>75</td>
<td>83</td>
<td>77</td>
<td>88</td>
<td>93</td>
<td>64</td>
<td>82</td>
<td>78</td>
<td>55</td>
<td>87</td>
<td>89</td>
<td>51</td>
</tr>
<tr>
<td>12</td>
<td>54</td>
<td>42</td>
<td>38</td>
<td>61</td>
<td>60</td>
<td>38</td>
<td>54</td>
<td>63</td>
<td>21</td>
<td>68</td>
<td>62</td>
<td>34</td>
</tr>
<tr>
<td>13</td>
<td>69</td>
<td>83</td>
<td>73</td>
<td>85</td>
<td>78</td>
<td>59</td>
<td>80</td>
<td>83</td>
<td>38</td>
<td>85</td>
<td>85</td>
<td>51</td>
</tr>
<tr>
<td>14</td>
<td>73</td>
<td>79</td>
<td>81</td>
<td>90</td>
<td>93</td>
<td>66</td>
<td>85</td>
<td>82</td>
<td>51</td>
<td>84</td>
<td>83</td>
<td>67</td>
</tr>
<tr>
<td>15</td>
<td>71</td>
<td>64</td>
<td>66</td>
<td>76</td>
<td>70</td>
<td>45</td>
<td>66</td>
<td>76</td>
<td>28</td>
<td>91</td>
<td>74</td>
<td>52</td>
</tr>
<tr>
<td>16</td>
<td>80</td>
<td>82</td>
<td>68</td>
<td>83</td>
<td>80</td>
<td>69</td>
<td>88</td>
<td>96</td>
<td>44</td>
<td>90</td>
<td>89</td>
<td>76</td>
</tr>
<tr>
<td>17</td>
<td>84</td>
<td>64</td>
<td>62</td>
<td>77</td>
<td>68</td>
<td>60</td>
<td>91</td>
<td>88</td>
<td>32</td>
<td>84</td>
<td>84</td>
<td>79</td>
</tr>
<tr>
<td>18</td>
<td>82</td>
<td>90</td>
<td>82</td>
<td>87</td>
<td>88</td>
<td>73</td>
<td>80</td>
<td>84</td>
<td>69</td>
<td>87</td>
<td>88</td>
<td>88</td>
</tr>
<tr>
<td>19</td>
<td>59</td>
<td>60</td>
<td>55</td>
<td>71</td>
<td>67</td>
<td>45</td>
<td>60</td>
<td>73</td>
<td>28</td>
<td>70</td>
<td>78</td>
<td>55</td>
</tr>
<tr>
<td>20</td>
<td>69</td>
<td>67</td>
<td>72</td>
<td>79</td>
<td>82</td>
<td>58</td>
<td>73</td>
<td>75</td>
<td>45</td>
<td>88</td>
<td>79</td>
<td>49</td>
</tr>
</tbody>
</table>

Note: Judges 1-10 = Orthographic Transcription Method; Judges 11-20 = Check-Slash Transcription Method
<table>
<thead>
<tr>
<th>J</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S5</th>
<th>S6</th>
<th>S7</th>
<th>S8</th>
<th>S9</th>
<th>S10</th>
<th>S11</th>
<th>S12</th>
</tr>
</thead>
<tbody>
<tr>
<td>J1</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>J2</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>J3</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>J4</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>J5</td>
<td>13.07</td>
<td>8.40</td>
<td>11.37</td>
<td>8.27</td>
<td>8.67</td>
<td>8.93</td>
<td>7.85</td>
<td>7.77</td>
<td>6.67</td>
<td>8.87</td>
<td>5.83</td>
<td>6.26</td>
</tr>
<tr>
<td>J6</td>
<td>8.17</td>
<td>7.08</td>
<td>6.37</td>
<td>5.50</td>
<td>5.37</td>
<td>5.65</td>
<td>5.50</td>
<td>5.87</td>
<td>5.68</td>
<td>6.95</td>
<td>5.57</td>
<td>5.25</td>
</tr>
<tr>
<td>J7</td>
<td>13.02</td>
<td>12.43</td>
<td>13.18</td>
<td>14.43</td>
<td>8.70</td>
<td>12.57</td>
<td>10.77</td>
<td>10.68</td>
<td>10.33</td>
<td>9.47</td>
<td>8.27</td>
<td>8.85</td>
</tr>
<tr>
<td>J9</td>
<td>23.77</td>
<td>15.20</td>
<td>17.93</td>
<td>15.53</td>
<td>9.93</td>
<td>15.28</td>
<td>15.38</td>
<td>10.43</td>
<td>20.23</td>
<td>11.90</td>
<td>10.38</td>
<td>15.78</td>
</tr>
<tr>
<td>J10</td>
<td>13.17</td>
<td>11.67</td>
<td>10.17</td>
<td>8.58</td>
<td>8.83</td>
<td>12.50</td>
<td>7.00</td>
<td>3.08</td>
<td>12.33</td>
<td>7.00</td>
<td>6.58</td>
<td>8.00</td>
</tr>
<tr>
<td>J11</td>
<td>4.42</td>
<td>6.52</td>
<td>5.15</td>
<td>4.08</td>
<td>3.32</td>
<td>4.77</td>
<td>4.30</td>
<td>5.20</td>
<td>6.75</td>
<td>6.73</td>
<td>3.33</td>
<td>4.67</td>
</tr>
<tr>
<td>J12</td>
<td>8.37</td>
<td>4.33</td>
<td>4.88</td>
<td>3.60</td>
<td>2.83</td>
<td>3.98</td>
<td>2.57</td>
<td>3.52</td>
<td>3.22</td>
<td>5.27</td>
<td>2.75</td>
<td>4.17</td>
</tr>
<tr>
<td>J13</td>
<td>5.07</td>
<td>4.13</td>
<td>4.85</td>
<td>3.67</td>
<td>4.02</td>
<td>4.97</td>
<td>3.53</td>
<td>3.25</td>
<td>3.60</td>
<td>5.15</td>
<td>2.97</td>
<td>4.17</td>
</tr>
<tr>
<td>J14</td>
<td>11.52</td>
<td>9.23</td>
<td>4.78</td>
<td>3.92</td>
<td>3.57</td>
<td>4.30</td>
<td>5.18</td>
<td>4.40</td>
<td>5.58</td>
<td>5.63</td>
<td>3.45</td>
<td>5.85</td>
</tr>
<tr>
<td>J16</td>
<td>12.42</td>
<td>13.15</td>
<td>11.38</td>
<td>7.72</td>
<td>8.17</td>
<td>8.58</td>
<td>6.35</td>
<td>7.93</td>
<td>7.82</td>
<td>6.47</td>
<td>7.35</td>
<td>5.87</td>
</tr>
<tr>
<td>J17</td>
<td>9.23</td>
<td>8.60</td>
<td>8.62</td>
<td>6.65</td>
<td>6.68</td>
<td>8.35</td>
<td>10.25</td>
<td>7.50</td>
<td>10.50</td>
<td>11.12</td>
<td>5.98</td>
<td>6.70</td>
</tr>
<tr>
<td>J18</td>
<td>4.33</td>
<td>4.67</td>
<td>5.15</td>
<td>3.67</td>
<td>4.98</td>
<td>6.73</td>
<td>4.05</td>
<td>3.65</td>
<td>5.93</td>
<td>6.95</td>
<td>2.87</td>
<td>3.09</td>
</tr>
<tr>
<td>J19</td>
<td>4.50</td>
<td>4.67</td>
<td>5.03</td>
<td>5.17</td>
<td>4.43</td>
<td>3.67</td>
<td>3.42</td>
<td>3.08</td>
<td>3.20</td>
<td>5.00</td>
<td>2.55</td>
<td>3.58</td>
</tr>
<tr>
<td>J20</td>
<td>8.92</td>
<td>8.70</td>
<td>8.77</td>
<td>6.30</td>
<td>5.07</td>
<td>6.12</td>
<td>7.45</td>
<td>6.17</td>
<td>7.92</td>
<td>5.13</td>
<td>5.72</td>
<td>6.78</td>
</tr>
</tbody>
</table>

Note: J = Judges; J1-10 = Orthographic Transcription Method; J11-20 = Check-Slash Transcription Method; ** indicates no time was available for this listener; S = Sample