Portland State University

PDXScholar

5-12-1997

Comparison of Intelligibility Estimation and Orthographic Transcription Methods by Preprofessional Speech-Language Pathologists

Kristi M. Mowe
Portland State University

Recommended Citation

THESIS APPROVAL

The abstract and thesis of Kristi M. Mowe for the Master of

Science in Speech Communication: Speech and Hearing

Sciences were presented May 12, 1997, and accepted by the

thesis committee and the department.

COMMITTEE APPROVALS:

Mary Gordon-Brannan, Chair

Doug Martin

Ellen Reuler

Nancy Benson
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 *24 June 1997*

ABSTRACT

An abstract of the thesis of Kristi M. Mowe for the Master
of Science in Speech Communication: Speech and Hearing
Sciences presented May 12, 1997.

Title: Comparison of Intelligibility Estimation and
         Orthographic Transcription Methods by
         Preprofessional Speech-Language Pathologists

When the fundamental means of communications is
speech, the main component for successful communication is
intelligibility. The speech of children with disordered
phonologies is often unintelligible; therefore, accurate and
reliable methods of measuring intelligibility are essential
when determining eligibility, and in selecting and providing
appropriate treatment.

The purpose of this study was to investigate the
relationship between the percentage estimation method and
the orthographic transcription method when measuring speech
intelligibility by preprofessional listeners. For this
study, the standard measurement of intelligibility was
defined as the percentage of words understood in a
continuous speech sample derived from orthographic
transcription of the sample.

Twelve speech samples were chosen from the original group of speakers in the Gordon-Brannan (1993) study according to varying levels of intelligibility. Three speakers were chosen from each of the four levels of phonological proficiency: adult-like, mild, moderate, and severe. Four graduate students in the Speech and Hearing Sciences Program at Portland State University served as preprofessional listeners. These listeners rated the children's percentage of intelligibility from continuous speech samples using listener estimation and orthographic transcription methods.

The two methods of measuring speech intelligibility investigated in this study were found to be strongly correlated overall among the 4 listeners, although the ranges among the listeners scores were wide for some of the speakers. Even though the correlation was strong between the two methods, there were substantial differences between the actual percentages derived from orthographic transcription and those derived from estimation for some listeners.

COMPARISON OF INTELLIGIBILITY ESTIMATION

AND ORTHOGRAPHIC TRANSCRIPTION METHODS

BY PREPROFESSIONAL SPEECH-LANGUAGE PATHOLOGISTS

by

KRISTI M. MOWE

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
1997

ACKNOWLEDGEMENTS

With many thanks

To Mary Gordon-Brannan, a superb advisor and friend,

To my mom, Ione Willson, for her love and support,

To my dad, Vern Willson, whose spirit lives on,

To my kids, Sam, Hadley, and Jackson, for their
    patience and understanding,

And to my husband, Will, for his encouragement and
    sticking with me through the years.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION AND STATEMENT OF PURPOSE

Introduction

Intelligibility has been variously defined as "the degree to which the speaker's intended message is recovered by the listener" (Kent, Weismer, Kent, & Rosenbek, 1989, p. 483), as the accuracy with which a message is conveyed and the listener's response to that message (Yorkston & Beukelman, 1980), and "the match between the intention of the speaker and the response of the listener" (Schiavetti, 1992, p. 13). In each of these perspectives, the role of the listener is given equal importance to that of the speaker. Intelligible speech is essential for effective communication to occur between people. Many researchers consider the measure of intelligibility to be the most important factor in determining communication competency (Bernthal & Bankson, 1988; Metz, Samar, Schiavetti, Sitler, & Whitehead, 1985; Subtelny, 1977); however, the question of how best to measure intelligibility remains an issue.

Speech-language pathologists (SLPs) use speech intelligibility measurements as one guideline for assessing the severity of a speech disorder. Van Riper and Emerick (1990) supported the notion of determining severity by stating "speech is defective when it is conspicuous, unintelligible, or unpleasant" (p. 34). Despite its

central role in evaluating speech disorders, intelligibility testing tends to be one of the most variable components in assessment protocols (Kent et al., 1989).

Since intelligibility is used, in part, to determine if speech intervention is necessary, accurate assessment is important. Intelligibility testing may also assist in ascertaining the type of intervention warranted and post-testing of intelligibility can help SLPs judge the efficacy of treatment (Weiss, 1982). Given the significance of intelligibility in assessing communicative competency and effectiveness in treating clients, it is important for SLPs to employ accurate, reliable, and valid measures in evaluating intelligibility.

There are various types of intelligibility measurements that clinicians and researchers currently employ. Objective measurements include word, sentence, and connected speech indentification procedures through orthographic transcription. Subjective measurements include percentage estimation as well as scaling procedures, such as interval scaling and direct magnitude estimation.

Interval scaling requires the listener to assign a number along a scale to estimate the degree of intelligibility of the speaker. Direct magnitude estimation involves rating intelligibility in comparison

with a representative speech sample. With orthographic transcription, a SLP writes down each word uttered in a connected speech sample, in sentences, or in word lists. For example, if the sample had 100 words transcribed and only 70 were understandable, then the child's speech would be judged to be 70% intelligible.

The most common method currently used by SLPs to assess speech intelligibility is to provide a subjective estimate of the percentage of words understood of a spontaneous speech sample of the child's speech (Gordon-Brannan, 1993). For example, an SLP may judge "Johnny was approximately 70% intelligible during conversation when the context was known and approximately 30% when the context was unknown."

Since the SLP's judgment is subjective, it is important to determine how closely the SLP's rating correlates to some of the more quantifiable measures such as orthographic transcription (Gordon-Brannan, 1993). Other investigations have explored the accuracy of estimates by practicing
SLPs (Clarke, 1997; Sugarman, 1994). What remains in question is how closely the preprofessional's percentage estimation of intelligibility correlates with orthographic transcription.

Statement of Purpose

The primary purpose of this study was to examine the relationship between the orthographic transcription method and the percentage estimation method when measuring speech intelligibility by preprofessional listeners. A secondary goal was to ascertain listener skill regarding the accuracy of percentage estimation as compared to the percentage derived from orthograhic transcription.

The following research questions were addressed:

1. What is the correlation between preprofessional listeners' estimations of the percentage of words understood and the percentage of words understood by the same listeners using the orthographic transcription method?

2. What is the range of variation between the two methods for individual listeners?

Definition of Terms

The following terms were defined as follows for this study:

AAPS - Arizona Articulation Proficiency Scale (Fudala & Reynolds, 1986).

ACI - Articulation Competence Index (Shriberg, 1993).

Assimilation - Influence of one sound in a word or phrase upon another sound to make it the same as the influencing sound in one or more features (i.e., voicing, place, or manner) (Hodson & Paden, 1991).

Backing - Replacing an anterior consonant with a posterior one (Hodson & Paden, 1991).

Cluster Reduction - One or more of the consonants of a consonant cluster is omitted (Hodson & Paden, 1991).

Consonant Sequence Omission - The omission of one or more sound segments from two or more contiguous consonants(Hodson & Paden, 1991).

DME - Direct magnitude estimation (subjective intelligibility measure) (Schiavetti, 1992).

Dysarthria - Motor speech disorder caused by weakness, paralysis, slowness, incoordination, or sensory loss in the muscle groups responsible for speech (Brookshire, 1992).

Dysfluency - Occurs when the forward flow of speech is interrupted abnormally by repetitions or prolongations of a sound, syllable, or articulatory

posture, or by avoidance and struggle behaviors (Van Riper & Emerick, 1990).

Dyspraxia - Neuromuscular speech problem characterized by inabililty or difficulty in performing speech acts voluntarily (Weiss, Gordon, & Lillywhite, 1987).

EAI - Equal-appearing interval (subjective intelligibility measure) (Schiavetti, 1992).

Final Consonant Deletion - The final singleton consonant in a word is omitted (Hodson & Paden, 1991).

Glottal Replacement - Substituting a glottal stop for a consonant. A glottal stop is not a distinctive sound in most dialects of the English language (Hodson & Paden, 1991).

Initial Consonant Deletion - A singleton consonant in a syllable is omitted (Hodson & Paden, 1991).

Laryngectomy - The surgical removal of the larynx (Van Riper & Emerick, 1990).

NTID - National Technical Institute for the Deaf (Johnson, 1975).

Phonological Deviation - Same as phonological process.

Phonological Process - A regularly occurring deviation in an individual's utterances, usually one that simplifies an adult phonological pattern (Hodson & Paden, 1991).

Postvocalic Singleton Omission - Final singleton consonant in a word is omitted. Same as final consonant

deletion (Hodson & Paden, 1991).

Prevocalic Singleton Omission - A singleton consonant in a
syllable is omitted. Same as initial consonant
deletion (Hodson & Paden, 1991).

SLP - Speech-language pathologist.

Stridency Deletion - Omission of the strident feature by
substitution of a nonstrident sound or by totally
omitting the target strident sound (Hodson & Paden,
1991).

Stopping - Substitution of stops for nonstop consonants
(Hodson & Paden, 1991).

Velar Deviation - A velar (i.e., /k/, /g/, and / /) is
omitted entirely or is replaced by a nonvelar
(Hodson, 1986).

Weak Syllable Deletion - Omission of an unstressed syllable
in a multisyllabic word (Hodson & Paden, 1991).

CHAPTER II

REVIEW OF THE LITERATURE

Intelligibility has been defined as the recognition of speech signals and the correspondence between the intention of the speaker and the response of the listener (Schiavetti, 1992). For a speech signal to be intelligible, both a speaker and a listener must be involved. When the listener matches all the words that are intended by the speaker, speech intelligibility is perfect. When none of the words are matched, speech intelligibility is zero. If there is a distortion or interference with the signal, intelligibility of the message may be affected and may ultimately cause difficulties in communication.

Having intelligible speech is considered to be one of the most basic components necessary for effective, competent communication (Connolly, 1986; Subtelny, 1977). During a child's school years, good intelligibility can be a determinant of school placement and an indicator of linguistic competence.

Many factors can influence an individual's intelligibility level. Some determiners of intelligibility may be etiological factors such as hearing impairment, dysarthria, dyspraxia, laryngectomy, cleft palate, and severe dysfluency (Connolly, 1986). Other components that may influence intelligibility include the listener's

familiarity with the speaker, the experience of the listener, contextual cues, the nature of the speech sample, the transmission media, phonological proficiency, the number and types of speech sound errors, and the parameters of frequency, duration, and stress (Beukelman & Yorkston, 1980; Hodson & Paden, 1981; Kelly, Dancer, & Bradley, 1986; Merwin, Goldstein, & Rothman, 1985; Weiss, 1982).

Factors Influencing Intelligibility

Factors influencing intelligibility can be numerous and sometimes difficult to identify (Hoops & Curtis, 1971; Merwin et al., 1985; Shriberg & Kwiatkowski, 1982; Weiss & Lillywhite, 1981). Major factors include listener familiarity with the speaker, the social context, message content, quality of the acoustic signal, visual cues, and linguistic complexity (i.e., sentence structure and length of utterance). Other elements that have an impact upon intelligibility are speech characteristics, including word pronunciation, adventitious speech sounds, articulation errors, dysfluency, phrasing, rate, stress, loudness, pitch, and voice quality (Monsen, 1983; Parkhurst & Levitt, 1978). In addition, phonological deviations, also called phonological processes, have been found to affect intelligibility negatively. Phonological deviations refer to patterns of speech sound errors rather than single phoneme errors. According to Hodson and Paden (1981), 11 phonological processes most highly influence the degree of

intelligibility. The phonological deviations that have the highest correlation with unintelligibility are cluster reduction, stridency deletion, stopping, liquid deviation, and assimilation. The other processes that influence intelligibility to a lesser degree are velar deviation, final consonant deletion, weak syllable deletion, prevocalic voicing, glottal replacement, and backing. Billman (1986) identified backing and prevocalic singleton omission as having the most detrimental effect on intelligibility.

Speech Intelligibility Assessment Procedures

Currently, two general measurement types are commonly used in research and clinical settings for the purpose of speech intelligibility assessment. Objective methods rely on computing an actual percentage of words understood in a speech sample. Subjective methods are impressionistic and involve the perception of a listener.

Objective Assessment Procedures

Objective speech intelligibility measurements involve a procedure wherein the listener writes down what the speaker says (Schiavetti, 1992). Cullinan, Brown, and Blalock (1986) stated that the measure of intelligibility usually employed in research has been the percentage of test items (sounds, words, or sentences) correctly transcribed or identified by a group of listeners. Methods of eliciting speech is often through word repetition, picture

identification tasks, and conversational speech samples
(Kent, Miolo, & Blodel, 1994; Morris, Wilcox, & Schooling,
1995; Weston & Shriberg, 1992).

Through word identification tests, intelligibility is
measured by calculating the percentage of words understood.
Weiss (1982) designed the Weiss Intelligibility Test in
which the listener transcribes 25 isolated words and
derives a percentage of intelligibility. A continuous
speech sample of 200 words is also transcribed by the
listener. The number of words understood and those not
understood is calculated. The total number of words
understood from the two sections are averaged together to
ascertain an overall intelligibility score. The score is
compared with intelligibility norms and is given a severity
rating level of normal, mild, moderate, severe, or
profound.

In the Children's Speech Intelligibility Test (CSIT)
(Kent et al., 1994), single words are elicited through a
word imitation task. The listener writes down the word or
chooses a word from a multiple-choice list of single words.
The listener's responses yield an overall percentage of
words understood as well as phonological contrast scores.

Another test for measuring percentage of
intelligibility is the Preschool Speech Intelligibility
Measure (P-SIM; Morris et al., 1995). In the P-SIM, the
child repeats a set of 50 randomly selected target words

after a model by the examiner. A listener identifies the tape-recorded responses and a percentage of correctly identified words is converted to a percentage score of intelligibility.

Monsen (1981) described a forced-choice test called the CID Word SPINE test (SPeech INtelligibility Evaluation). The test was designed to measure the speech intelligibility of children with severe to profound hearing impairments. Monsen's SPINE test involves a write-down procedure that measures the phonological integrity of a child's speech. In the SPINE, the child is familiarized with target words that are printed on cards. The child reads the words aloud and the examiner selects the word from an answer sheet. A percentage of words correctly understood is then calculated.

A traditional approach for measuring intelligibility is orthographic transcription of each word in a conversational speech sample. The percentage of intelligibility score is the number of words correctly understood in a 100-word speech sample (Kent et al., 1994).

Subjective Assessment Procedures

Rating scale procedures are subjective measures of intelligibility. The most commonly used methods are equal-appearing interval scaling and direct magnitude estimation (Darley, Aronson, & Brown, 1969; Schiavetti, 1992). With these methods, a number is assigned by a listener to a

speaker's speech sample. This number reflects the listener's perception of the speaker's overall intelligibility.

In equal-appearing scaling techniques, a number is placed on a continuum that represents the listener's perception of a speaker's intelligibility. The continuum is generally a scale numbered from 1 to 5, 7, or 9, with the numbers representing the degree of intelligibility, (e.g., 1 representing completely unintelligible to 5, representing completely intelligible). Some rating scales are labeled with descriptors. For example in the National Technical Institute for the Deaf scale (Johnson, 1975) the following descriptors correspond with the numbered rating scale:

1. Speech cannot be understood.

2. Speech is very difficult to understand--only isolated words and phrases are intelligible.

3. Speech is difficult to understand; however, the gist of the content can be understood.

4. Speech is intelligible with the exception of a few words or phrases.

5. Speech is completely intelligible. Direct magnitude estimation (DME) does not require the listener's rating to fit upon a finite scaled continuum. Instead, the listener judges each speech sample with a number that is proportional to the speech intelligibility among many

speech samples of the speaker (Schiavetti, 1992). The DME can be used with or without a standard/modulus. A speech sample is collected and given a modulus or subjective value of intelligibility by the listener. The remaining speech samples are judged for intelligibility against this modulus number. The listener assigns any number to the first speech sample when a standard and modulus are not used. All subsequent speech samples are assigned numbers that relate to the perceived amount of intelligibility of the various speech samples in comparison with the first sample (Schiavetti, 1992).

Objective Versus Subjective Assessment Procedures

Given the different methods of assessing speech intelligibility, the SLP must decide which is the most appropriate way to measure a client's intelligibility. The method chosen depends almost entirely on the purpose for which the testing is being done. When the clinician wants to know what types of errors are made, an articulation and/or phonological test plus a percentage of words understood may be warranted. If the clinician wants to know how well speech is understood, the speech sample analyzed should be highly representative of everyday speech (Speaks, Parker, Harris, & Kuhl, 1972).

There may not be a single assessment method that is appropriate and/or completely accurate for evaluating children's speech intelligibility. Both objective and

subjective measurements have advantages and disadvantages.
Many clinicians prefer to use single-word intelligibility
tests because they are easier and less time consuming to
administer and score. Also, word lists can be reordered
for repeated measurement. Many researchers have stated
that word identification measures are much more
advantageous as the results are easier to use, criterion
validity is high, and a description of the disorder is
provided (Beukelman & Yorkston, 1979; Metz, Samar,
Schiavetti, Sitler, & Whitehead, 1985; Samar & Metz, 1988;
Schiavetti, 1992; Yorkston & Beukelman, 1978).

Some researchers have indicated that contextual
intelligibility scores are higher than isolated word
intelligibility scores (Goetzinger, 1978; Subtelny, 1977).
Many studies have indicated that intelligibility scores for
individual words are not good predictors of intelligibility
scores for continuous speech (Clark & Stemple, 1982; Cox &
McDaniel, 1984; Cullinan et al., 1986). Sitler,
Schiavetti, and Metz (1983) have suggested that the
contextual intelligibility measure is more appropriate
because speech is normally produced in a context with
linguistic and acoustic redundancies that are lacking in
speech produced as isolated words. Kelly et al. (1986)
concurred that "for normal speech the redundancy to the
context is usually unimportant, but for deviant speech it
can be essential to its comprehension." It should be noted

that the words and/or sentences in an objective test are not produced spontaneously, nor within the context of a dialogue. Additionally, Boothroyd (1985) stated that the use of ready-made sentences does not take linguistic competence into account.

Scaling procedures are considered advantageous by some because they are simple to use, expensive equipment is not required, and they can be applied in almost any situation (Metz, Schiavetti, & Sitler, 1980). However, a problem relating to subjective measurement is the limited diagnostic and prescriptive value of the results. A percentage score or a placement on a rating scale may indicate the need for speech remediation, but it does not provide analytic detail or any answers as to why there is a problem.

SLPs typically rely on estimates of the percentage of intelligible words in a continuous speech sample (Shriberg & Kwiatkowski, 1982). Shriberg and Kwiatkowski stated that tracking the number of sound errors does not adequately assess intelligibility because the same pattern of errors becomes more intelligible as the listener becomes familiar with the pattern.

Summary and Implications

Speech intelligibility appears to be one of the primary indicators of communicative competency. Since communicative competency and effectiveness is a major goal in treating

clients who are speech delayed/disordered, it is important for clinicians to employ reliable, valid measures when evaluating intelligibility.

Two distinct types of methods are currently used to measure intelligibility: objective and subjective. There are advantages and disadvantages to both. A combination of the two may be the best solution. Given time constraints in the diagnostic session, the quickest method of evaluating speech intelligibility may often be used rather than the most reliable or valid method. A relatively quick and reliable method of assessing intelligibility needs to be devised. In this study, listener judgments of continuous speech samples versus orthographic transcription were explored in order to gain information in creating a fast, accurate method of evaluating speech intelligibility.

CHAPTER III

METHOD

The purpose of this study was to examine the
correlation between preprofessional listeners' estimation of
speech intelligibility and the orthographic transcription
method of measuring speech intelligibility. Twelve speech
samples were analyzed to determine percentage of
intelligibility using estimation and orthographic
transcription.

In a study by Gordon-Brannan (1993), four graduate
students at Portland State University estimated and
orthographically transcribed 48 speech samples via
audiotape. An objective intelligibility percentage was
calculated for each transcribed sample. In this study,
preprofessional listeners provided a percentage estimate of
intelligibility and orthographically transcribed 12 speech
samples via audiotape. Objective percentages were
determined for each transcribed sample.

Participants

Speakers

The speakers were selected from the greater Portland,
Oregon metropolitan area preschools and speech-language
pathology caseloads by Gordon-Brannan (1993) in connection
with her doctoral dissertation. Informed consent forms were

signed by their parents before any hearing, speech, and receptive language testing was done with their children. The parents also filled out a questionnaire on their child's speech, hearing, developmental history, and socioeconomic level.

In the Gordon-Brannan (1993) study, 48 children served as speakers who ranged phonologically from error-free speech to numerous omission errors. They ranged in age from 4:0 to 5:6. These children were selected according to testing results and the following criteria:

1. No known neurological, motor, or physical impairment.

2. No noticeable upper respiratory congestion or voice quality disorder.

3. Age-appropriate receptive vocabulary.

4. Hearing sensitivity within the normal/mild loss range.

The speakers who qualified for inclusion in the study were placed into 4 subgroups on the basis of their phonological proficiency/deficiency. The four categories of speech characteristics were labeled adult-like, mildly deficient, moderately deficient, and severely deficient; thus a wide range of severity was represented.

Listeners

For this study, the group of listeners was comprised of 4 graduate students who were selected from the Speech and Hearing Sciences Program at Portland State University, and who had between 200 and 250 hours of clinical experience. Additional criteria for the listeners were that they were unfamiliar with the speakers and reported normal hearing. Subjects were recruited via fliers and announcements in graduate level classes in the Speech and Hearing Sciences program. The 4 listeners signed informed consent forms prior to estimating and transcribing the 12 speech samples (Appendix A).

<div align="center">Procedures</div>

Speech Sample Collection

The speech samples used for this study were originally created for Gordon-Brannan's (1993) study on intelligibility. The tapes consisted of connected speech samples that were elicited from each speaker in a sound-treated recording room at Portland State University. Each speaker was both audiotaped and videotaped. The equipment used to audio and videotape the samples included a Panasonic camcorder, VHS Brand Recorder, and a Sharp SX D2000 digital audiotape recorder. An AKG Model C451 microphone was used to record the speech samples.

The speech samples were collected by the investigator and each sample contained 100 words in length that were evoked through the use of books, pictures, and a set of open-ended questions. The utterances of both the investigator and the speaker were audio and videotaped. The speech samples were dubbed onto listener tapes. The speech samples were then placed in random speaker order and were used for later transcription and rating.

Estimation and Transcription Procedure

In the original study, transcription scoring keys for each continuous speech sample were prepared from the orthographic transcriptions for the investigator to verify the child's words. The parent(s) of the children in the moderate to severe deficiency groups reviewed the content of the transcription and verified or corrected the listener's interpretation of their child's speech.

Of the original 48 100-word connected speech samples, 12 speech samples were used in this study. These samples had been elicited from 4 girls and 8 boys between the ages of 4:1 and 5:1. To obtain these 12 samples, three were chosen at random from each severity rating category: adult-like, mildly deficient, moderately deficient, and severely deficient.

For this study, the listeners estimated an overall percentage of words understood for the 12 selected speech

samples after listening to them via a digital audiotape. The listeners participated in a brief training procedure at the beginning of the session. During training, listeners were given a score sheet, oral and written instructions, and answers to any questions (Appendix B).

The 12 speech samples were dubbed onto analogue audiotapes from the orignal digital audiotapes. After a two-week time period, the listeners listened to the same speech samples on their own analogue tape recorders and orthographically transcribed each speaker's connected speech sample. Written instructions were given to each listener for this task (Appendix C).

Scoring

In the original Gordon-Brannan (1993) study, the listeners' orthographic transcriptions were compared with the transcription keys compiled by the investigator. The listener's transcribed word was considered incorrect if the listener identified a word that differed from the prepared transcription key. Morphological form differences were not considered incorrect. The percentage of words understood by each listener for each continuous speech sample was computed.

For this study, the subjects wrote down word-for-word what they understood the child to say. Words that were not understood were denoted by using an "X" as a place marker.

Only complete words were counted as intelligible. The listeners' orthographic transcriptions were compared to the transcription keys from the original study. The same rules were applied regarding incorrect word identification and morphological form differences. Percentage of words understood was then derived from each transcribed speech sample.

Data Analysis

A Pearson product-moment correlation (Pearson r) was used to address the research question of the correlation between the estimation method for measuring speech intelligibility and the method of orthographic transcription for preprofessional listeners. A Pearson r was also used to determine interjudge reliability between each pair of graduate student listeners for each method for a total of 12 correlations. Finally, the data were further analyzed by visual inspection to compare the actual percentages derived from orthographic transcription with the estimated percentage.

CHAPTER IV

RESULTS AND DISCUSSION

Results

The purpose of this study was to examine the relationship between the orthographic transcription method and the percentage estimation method when measuring speech intelligibility. Specifically, the focus of this research project was to determine the correlation between preprofessional listeners' estimations of the percentage of words understood and the percentage of words understood by the same listeners using the orthographic transcription method. A secondary goal was to ascertain listener skill regarding the accuracy of percentage estimation as compared to the percentage derived from orthographic transcription. The results are presented in the following sections. Raw data appear in Appendix D.

Reliability Analysis

Preliminary to addressing the research questions, interjudge reliability of the listeners was examined by applying Pearson r. The percentage data for each speaker sample by each listener appear in Appendixes E & F. Tables 1 and 2 provide the correlation coefficient data for interrater reliability for the estimation and orthographic transcription methods, respectively. Each listener's scores were correlated with the other listeners for both measures. All correlations were statistically significant at or beyond

Table 1

<u>Correlation Matrix for Comparison of Listeners by Pairs for</u>
<u>the Estimation Method</u>

Listener	1	2	3	4
1	1.00			
2	.94*	1.00		
3	.90*	.84*	1.00	
4	.97*	.91*	.89*	1.00

*p ≤ .05

Table 2

<u>Correlation Matrix for Comparison of Listeners by Pairs for</u>
<u>the Orthographic Method</u>

Listener	1	2	3	4
1	1.00			
2	.93*	1.00		
3	.96*	.94*	1.00	
4	.94*	.90*	.97*	1.00

*p ≤ .05

the .05 level. Pearson-r correlations for the percentage-
of-words understood in continuous speech ranged from .84 to
.97 for the six pairs of listeners using the estimation
method, and from .90 to .97 for the six pairs of listeners
using the orthographic transcription method. These
correlations indicate the 4 listeners were in slightly
higher agreement using the orthographic transcription method
and in general agreement overall.

The relationship among the listeners was examined
further by computing the range of percentage scores derived
by the two methods for each speaker. Review of the data
revealed a range of 0 (speaker 8) to 56 (speaker 20), with a
mean range of 22, for the estimation method. For the
orthographic method, less variability was noted with a range
of 3 (speaker 5) to 27 (speaker 25), with a mean range of
16. This would suggest that the orthographic values are
more consistent among listeners than the estimation values.

Research Question

The first research question investigated was: What is
the correlation between preprofessional listeners'
estimations of the percentage of words understood and the
percentage of words understood by the same listeners using
the orthographic transcription method? The mean percentage
of intelligibility for each speaker for both measures is
provided in Appendix G. A Pearson r was used to determine

the correlation between the two measures for each listener. The Pearson-r correlations ranged from .88 to .98 (p ≤ .05). These results indicate speech intelligibility derived by estimation is highly correlated with the percentage of intelligibility dervied from orthographic transcription. To provide additional information regarding the relationship between the two methods, Pearson-r correlations were calculated using the data from this study and the original study (Gordon-Brannan, 1993). The correlations coefficients ranged from .88 to .98 (p ≤ .05) for the estimation method and .93 to .98 (p ≤ .05) for the orthographic method.

To address the range of variation between the two methods for individual listeners, raw percentage scores for the estimation method revealed a range of 0 (speaker 8) to 56 (speaker 20), with a mean range of 21. A range of 3 (speaker 5) to 27 (speaker 25) was noted for the orthographic method, with a mean range score of 13. Again, this shows less variability with the orthographic transcription method, suggesting more objectivity and consistency.

These results indicate speech intelligibility derived by listener estimation is highly correlated with the percentage of intelligibility derived from orthographic transcription. These values also indicate that within each method there was a strong linear association between the

listeners across the 12 samples. This would imply that the two methods could be used interchangeably. However, analysis of the data demonstrated some variability of scores between the two methods for given listeners. In this study, the orthographic transcription scores generally produced a more accurate picture of percentage of intelligibility for each speaker when compared to the scores from the original study (Gordon-Brannan, 1993). The estimation scores in this study were not as reliable or accurate in determining percentage scores for intelligibility when compared to the original study scores.

Discussion

Two methods for measuring speech intelligibility, estimation and orthographic transcription, were investigated in this study. The orthographic transcription method is considered to be the standard measurement by many speech-language pathologists because it involves writing down each word the speaker utters in a continuous speech sample (Schiavetti 1992); however, the subjective method of listener estimation appears to be more widely employed probably because of time constraints and the lack of need to tape record the client's utterances (Shriberg & Kwiatkowski, 1982). In this study, listener estimation and orthographic transcription methods for measuring speech intelligibility

were compared to determine the correlation between the two measures. The results indicated the two measures were highly correlated, although actual percentages between listeners for the same speaker widely varied.

The preprofessional listeners demonstrated high interjudge reliability, even though some speakers received a wide range of estimated percentages from the 4 listeners (Appendix E). One speaker (speaker 8) was judged identically by all 4 listeners as 98% intelligible. At the other end of the spectrum, another speaker (speaker 20) was judged 60% intelligible by Listener 1, but only 4% intelligible by Listener 4, which represents a range of 56%. Notable differences among listener estimations (over 20%) occurred for 5 of the remaining 12 samples. These results indicate that, even though the listeners were in overall general agreement with each other, their actual estimations for some individual speakers varied greatly. A mean range of 22 was calculated for the estimation method across the 4 listeners. Further inspection of the results shows that there was higher agreement between listener estimation percentages of speakers who were more intelligible. Conversely, the data also suggest that estimation agreements among listeners decreased with the speakers who were the least intelligible.

In utilizing the orthographic transcription method, the listeners were in general agreement with each other as well; however, the transcriptions of individual speakers reveal a wide range of percentages of words understood (Appendix F). None of the speakers were found to be 100% intelligible by the transcription method. Speaker 25 received the largest percentage difference (27%), with Listener 1 transcribing 89% words correctly and Listener 2 transcribing 62% words correctly. Even though differences of 20% or more occurred in 5 of the 12 transcribed samples, the mean range difference for the orthographic method (mean = 16) was less than the estimation method (mean = 22).

Since a listener is necessary to determine the intelligibility of a speech signal, listener experience and behavior are very important factors. The influence of listener experience in assessing speech intelligibility has been documented in several studies (Markides, 1970; McGarr, 1981; Monsen, 1978). With both methods, listener differences are most likely related to the listener's ability to identify individual phonemes and assign meaning. Each listener possesses varying amounts of experience and skill. The results of this study indicate that the listeners who participated had varying levels of perceptual skill, demonstrated by individual listener data.

Listener Estimations versus Orthographic Transcription

The listener estimation method of assessing intelligibility is entirely subjective; therefore, accuracy and reliability of the method are often in question (Kent et al., 1994). On the other hand, even though the task of transcribing language samples is probably the most objective method of assessing intelligibility, it is still somewhat subjective in nature because of inevitable individual and interrater differences. Across the listener group, results of this study indicated there is a high correlation between the estimation and orthographic measures. At the individual listener level, however, data indicated that the methods do not necessarily yield equivalent percentages in measuring intelligibility.

It is important to report that overall the listeners were fairly consistent in their ability to rate speech samples, as shown by the strong Pearson-r scores; however, in examining the raw data, individual listeners' percentages varied widely. Intelligibility scores derived from the estimation method ranged from 4% to 100%, and 24% to 99% for the orthographic transcription method. Listener 1 had 7 of the highest scores on the estimation percentages and also 7 of the highest orthographic scores among the 4 listeners. It should be noted that this listener reported spending the most time in transcribing the samples, thus suggesting more

accuracy with the orthographic transcription method when more time is taken. Listener 1 had estimation scores ranging from 60% to 98% (mean = 85.08) for the estimation method and scores ranging from 46% to 99% (mean = 76.5) for the orthographic method. The difference between the two methods for this listener ranged from 1 to 21 percentage points. The mean range difference in comparing both methods for this listener was 8.75.

Listener 2 estimated only 1 of the 12 samples the highest among the 4 listeners. With the orthographic transcription method, listener 2 had 10 of the lowest orthographic scores among the 4 listeners and reported the least amount of time in transcribing the samples. This suggests that if samples are hastily transcribed, a less accurate percentage may be obtained. Listener 2 had estimation scores that ranged from 40% to 100% (mean = 74) for the estimation method and a range of 32% to 93% (mean = 62.9) for the orthographic transcription method. The difference between the methods for this listener ranged from 3 to 24. The mean range difference in comparing both methods was the greatest of all 4 listeners at 13.08.

Listener 3 had one of the highest estimated percentages and one of the highest orthographic transcription scores among the 4 listeners. Estimates ranged from 20% to 100% (mean = 68.8) and the range was 43% to 96% (mean = 73.25)

for the orthographic method. The difference between the methods for this listener ranged from 0 to 23. The mean range difference in comparing the two methods for this listener was 10.25.

Listener 4 had estimation scores ranging from 4% to 100% (mean = 72.75) and a range of 24% to 94% (mean = 72.16) for the orthographic method. The difference between the two methods for this listener ranged from 0 to 20. The mean range difference of the two methods for listener 4 was 8.58, showing the greatest consistency between the two methods of the 4 listeners, although not notably different from Listener 1. Listener 4 had the lowest estimation percentage and the lowest orthographic score on one of the samples (speaker 20); however, it should be noted that this speaker had the lowest estimation percentages and orthographic transcription scores from all 4 listeners. This speaker was also the least intelligible in this study according to the original study (Gordon-Brannan, 1993).

Visual inspection of the raw data of "most intelligible" to "least intelligible" speakers appear to show that there is closer agreement and similar estimation ratings among all 4 listeners with the speakers of higher intelligiblity. An exception to this observation is listener 3, estimating speaker 25 at 60%, while Listeners 1 and 4 estimate this speaker at 95% intelligible, and

Listener 2 estimating 85%. Larger discrepancies with the estimation method appear among listeners as speaker intelligibility decreases (Appendix H). When examining "most intelligible to "least intelligible," no clear pattern was evident with the orthographic transcription method.

Components which influence intelligibility can be numerous and sometimes difficult to identify (Hoops & Curtis, 1971; Merwin et al., 1985; Shriberg & Kwiatkowski, 1982). Weiss (1982) listed 21 factors that affect intelligibility (Appendix I). Many factors may have affected the listeners' ability to understand the speech samples in this study. All of the 4 listeners were graduate students with 200 to 250 hours of clinical practice. Amount of experience working with speech and language disordered children and/or motivation may have influenced the accuracy and/or reliability of the two intelligibility measures.

One factor that may have increased the accuracy of the listeners' estimations and transcriptions was the examiner's utterances on the sample tapes. The speech samples included the conversation between the examiner and individual subjects. Even though the examiner attempted to control for this, two of the listeners reported that the content of what the child said was sometimes understood because of the examiner's utterances. Factors that may have hindered the listeners' understanding were fatigue and desensitization to

the stimulus material. Listening to children of varying phonological proficiency telling the same story may have been increasingly tiresome and difficult.

Even though the two measures are highly correlated, the time required to use each of the methods varied greatly. The listeners estimated the intelligibility of 12 speech samples in approximately 1 hour. The listeners spent a reported 3 to 10 hours orthographically transcribing the 12 samples. Three out of the 4 listeners reported the transcription task as "tedious" and/or "very difficult." These results indicate that the estimation method is obviously more time efficient.

The conclusion to be drawn from this study is that the listener estimation and the orthographic transcription methods of assessing intelligibility are highly correlated. However, given inevitable individual listener differences, the two methods should not automatically be considered equivalent in their application. The greater variability within the estimation method should be taken into account when deciding if this is a universally valid method to employ. Other factors that influence which method to use include speaker familiarity, content familiarity, speech severity of the speaker, time constraints, and listener skill. The results of this study indicate that the percentage of intelligibility determined by either method

would be highly dependent upon the perceptual skill of the listener.

CHAPTER V

SUMMARY AND IMPLICATIONS

Summary

When the fundamental means of communication is speech, the main component for successful communication is intelligibility. The speech of children with disordered phonologies is often unintelligible; therefore, accurate and reliable methods of measuring intelligibility are essential when determining eligibility, and in selecting and providing appropriate treatment.

The purpose of this study was to investigate the relationship between the percentage estimation method and the orthographic transcription method when measuring speech intelligibility by preprofessional listeners. For this study, the standard measurement of intelligibility was defined as the percentage of words understood in a continuous speech sample derived from orthographic transcription of the sample.

Twelve speech samples were chosen from the original group of speakers in the Gordon-Brannan (1993) study according to varying levels of intelligibility. Three speakers were chosen from each of the four levels of phonological proficiency: adult-like, mild, moderate, and severe. Four graduate students in the Speech and Hearing Sciences Program at Portland State University served as

preprofessional listeners. These listeners rated the children's percentage of intelligibility from continuous speech samples using listener estimation and orthographic transcription methods.

The two methods of measuring speech intelligibility investigated in this study were found to be strongly correlated overall among the 4 listeners, although the ranges among the listeners scores were wide for some of the speakers. Even though the correlation was strong between the two methods, there were substantial differences between the actual percentages derived from orthographic transcription and those derived from estimation for some listeners.

<div align="center">Implications</div>

<u>Clinical</u>

The mean score for listener estimation was 75%, and the mean score for orthographic transcription was 71%. The Pearson r correlation (.87) indicated the two measures are strongly correlated when rating speech intelligibility of young children.

Even though intelligibility is not the only component considered in determining eligibility and level/priority of service, it is an essential element. Statistical results of this study support the hypothesis that the listener estimation method is highly correlated to the orthographic

method when measuring speech intelligibility; however given the range and variability in the actual estimation percentages as compared to the orthographic transcription method, the two methods are not equivalent. Clinicians should be aware that differing methods for determining percentage of intelligibility yield differing results. There are inevitable individual differences when measuring intelligibility and the results of this study demonstrate how each clinician comes to the task with varying levels of skill in this area. Clinicians need to be aware of their own abilities to judge a speech sample accurately. SLPs could determine competency by testing their estimation percentages against their transcriptions for individual speakers. If the clinician finds the two methods to be consistently equal, either method could be employed with confidence. If a clinician consistently overestimates intelligibility, the child may be found ineligible for speech services and/or receive a lesser amount of service time.

For children eligible and receiving speech services, periodic evaluation of intelligibility is necessary to assess treatment procedures, progress, and level of service. Accurately assessing speech intelligibility is therefore an important factor in determining if treatment methods are effective. Again, clinicians need to be aware of their own

ability to judge a speech sample in deciding which method to employ.

Research

It is important for clinicians to be proficient in evaluating and administering intelligibility intruments since intelligibility measurement is a major factor in determining eligibility and level of service. This study focused on the listener estimation method commonly employed by many SLPs and its correlation to the orthographic transcription method. The results of this study at the group level give general validity to the listener estimation method. To further establish the validity of this method, future studies could increase sample size, with both speakers and 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 may show the effects of listener perception.

If the same speech samples are used in a future study, the examiner's voice could be deleted from the audiotape to obtain an intelligibility score purely derived from listening to the speaker. However, since some contextual cues are already provided, additional information could be provided through videotapes. It would also be interesting to investigate the correlation between other intelligibility measures, such as a tracking or rating method and listener

estimation.

The variability in estimation percentages at the individual listener level was wide in this study. Possibly the ranges of percent intelligibility would be narrower if experienced SLPs were the listeners rather than graduate students.

Clinicians need to be held accountable for providing the most accurate and reliable battery of tests and treatment to their clients. With further research into the subject of intelligibility measurements, clients can be assured of assessments that are valid, appropriate, and efficacious.

References

Bernthal, J. E., & Bankson, N. W. (1988). Articulation and phonological disorders. Englewood Cliffs, NJ: Prentice Hall.

Beukelman, D. R., & Yorkston, K. M. (1979). The relationship between information transfer and speech intelligibility of dysarthric speakers. Journal of Communication Disorders, 12, 189-196.

Beukelman, D. R., & Yorkston, K. M. (1980). Influence of passage familiarity on intelligibility estimates of dysarthric speech. Journal of Communication Disorders, 13, 33-41.

Billman, K. S. (1986). Phonological processes and intelligibility of spontaneous utterances in young children. Unpublished master's thesis, San Diego State University, San Diego, CA.

Boothroyd, A. (1985). Evaluation of speech production of the hearing impaired: Some benefits of forced-choice testing. Journal of Speech and Hearing Research, 28, 185-196.

Brookshire, R. H. (1992). An introduction to neurogenic communication disorders (4th ed.). St. Louis, MO: Mosby-Year Book, Inc.

Clark, J., & Stemple, J. (1982). Assessment of three modes of alaryngeal speech with a synthetic sentence

identification task (SSI) in varying message to competition ratios. Journal of Speech and Hearing Research, 25, 333-338.

Clarke, H. (1997). Gross estimation: A study of clinical validity for measuring intelligibility. Unpublished master's thesis, Portland State University, Portland, OR.

Connolly, J. H. (1986). Intelligibility: A linguistic view. British Journal of Disorders of Communication, 21, 371-376.

Cox, R. M., & McDaniel, D. M. (1984). Intelligibility ratings of continuous discourse: Application to hearing aid selection. Journal of Acoustical Society of America, 76, 758-766.

Cullinan, W. L., Brown, C. S., & Blalock, P. D. (1986). Ratings of esophageal and tracheoesophageal speech. Journal of Communication Disorders, 19, 185-195.

Darley, F., Aronson, A. E., & Brown, J. R. (1969). Differential diagnostic patterns of dysarthria. Journal of Speech and Hearing Research, 12, 246-269.

Fudala, J. B., & Reynolds, W. M. (1986). Arizona articulation proficiency scale (2nd ed.). Los Angeles: Western Psychological Services.

Goetzinger, C. P. (1978) Word discrimination testing. In J. Katz, (Ed.), Handbook of clinical audiology.

Baltimore: Williams & Wilkins.

Gordon-Brannan, M. (1993). Speech intelligibility assessment of young children with varying levels of phonological proficiency/deficiency. Unpublished doctoral dissertation, Wichita State University, Wichita, KS.

Hodson, B. W. (1986). <u>The assessment of phonological processes-revised</u>. Austin, TX: Pro-Ed.

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

Hodson, B. W., & Paden, E. P. (1991). <u>Targeting intelligible speech</u> (2nd ed.). Austin, TX: Pro-Ed.

Hoops, H. R., & Curtis, J. F. (1971). Intelligibility of the esophageal speaker. <u>Archives of Otolaryngology</u>, <u>93</u>, 300-303.

Johnson, D. D. (1975). Communication characteristics of NTID students. <u>Journal of the Academy of Rehabilitative Audiology</u>, <u>8</u>, 17-32.

Kelly, C., Dancer, J., & Bradley, R. (April, 1986). Correlation of SPINE test scores to judges' ratings of speech intelligibility in hearing-impaired children. <u>The Volta Review</u>, 145-150.

Kent, R. D., Miolo, G., & Blodel, S. (1994). Intelligibility of children's speech: A review of evaluation

procedures. American Journal of Speech-Language Pathology, 4, 81-95.

Kent, R. D., Weismer, G., Kent, J. F., & Rosenbek, J. C. (1989). Toward phonetic intelligibility testing in dysarthria. Journal of Speech and Hearing Disorders, 54, 482-499.

Markides, A. (1970). The speech of deaf and partially hearing children with special reference to factors affecting intelligibility. British Journal of Disorders of Communication, 5, 127-140.

McGarr, N. S. (1981). The effect of context on the intelligibility of hearing and deaf children's speech. Language and Speech, 24, 255-264.

Merwin, G. E., Goldstein, L. P., & Rothman, H. B. (1985). A comparison of speech using artificial larynx and tracheoesophageal puncture with valve in the same speaker. Laryngoscope, 95, 730-734.

Metz, D. E., Samar, V. J., Schiavetti, N., Sitler, R. W., & Whitehead, R. L. (1985). Acoustic dimensions of hearing-impaired speakers' intelligibility. Journal of Speech and Hearing Research, 28, 345-355.

Metz, D. E., Schiavetti, N., & Sitler, R. W. (1980). Toward an objective description of the dependent and independent variables associated with intelligibility assessments of hearing-impaired adults. In J.D. Subtelny,

Speech assessment and speech improvement for the hearing-impaired (pp.72-81). Washington D.C.: The Alexander Graham Bell Association for the Deaf, Inc.

Monsen, R. B. (1978). Toward measuring how well hearing-impaired children speak. Journal of Speech and Hearing Disorders, 21, 197-219.

Monsen, R. B. (1981). A usable test for the speech intelligibility of deaf talkers. American Annals of the Deaf, 126, 945-852.

Monsen, R. B. (1983). The oral speech intelligibility of hearing-impaired talkers. Journal of Speech and Hearing Disorders, 48, 286-296.

Morris, S. R., Wilcox, K. A., & Schooling, T. L. (1995). The preschool intelligibility measure. American Journal of Speech Language Pathology, 4, 24-31.

Parkhurst, B. G., & Levitt, H. (1978). The effect of selected prosodic errors on the intelligibility of deaf speech. Journal of Communication Disorders, 11, 249-256.

Samar, V. J., & Metz, D. E. (1988). Criterion validity of speech intelligibility rating-scale procedures for the hearing-impaired population. Journal of Speech and Hearing Research, 31, 307-316.

Schiavetti, N. (1992). Scaling procedures for the measurement of speech intelligibility. In R. D. Kent, Intelligibility in speech disorders: Theory, measurement

and management (pp. 11-34). Amsterdam and Philadephia: John Benjamins Publishing Company.

Shriberg, L. D. (1993). Four new speech and prosody measures for genetics research and other studies in developmental phonological disorders. Journal of Speech and Hearing Research, 36, 105-140.

Shriberg, L. D., & Kwiatkowski, J. (1982). Phonological disorders III: A procedure for assessing severity of involvement. Journal of Speech and Hearing Disorders, 47, 256-270.

Sitler, R. W., Schiavetti, N., & Metz, D. E. (1983). Contextual effects in the measurement of hearing-impaired speakers' intelligibility. Journal of Speech and Hearing Research, 26, 30-34.

Speaks, C., Parker, B., Harris, C., & Kuhl, P. (1972). Intelligibility of connected discourse. Journal of Speech and Hearing Research, 15, 590-602.

Subtelny, J. (1977). Assessment of speech with implications for training. In F. Bess (Ed.), Childhood deafness. New York: Grune & Stratton.

Sugarman, N. (1994). A comparison between trained ear estimation and orthogrphic transcription when measuring speech intelligibility of young children. Unpublished master's thesis, Portland State University, Portland, OR.

Van Riper, C., & Emerick, L. (1990). <u>Speech Correction</u>. (8th ed.). Englewood Cliffs, NJ: Prentice Hall.

Weiss, C.E. (1982). <u>Weiss intelligibility test</u>. Tigard, OR: CC Publications.

Weiss, C. E., Gordon, M. E., & Lillywhite, H.S. (1987). <u>Clinical management of articulatory and phonologic disorders</u>. Baltimore: Williams & Wilkins.

Weiss, C. E., & Lillywhite, H. S. (1981). <u>Communicative disorders</u> (2nd ed.). St. Louis: C.V. Mosby Company.

Weston, A. D., & Shriberg, L. D. (1992). Contextual and linguistic correlates of intelligibility in children with developmental phonological disorders. <u>Journal of Speech and Hearing Research</u>, <u>35</u>, 1316-1332.

Yorkston, K. M., & Beukelman, D. R. (1978). A comparison of techniques for measuring intelligibility of dysarthric speech. <u>Journal of Communication Disorders</u>, <u>11</u>, 499-512.

Yorkston, K. M., & Beukelman, D. R. (1980). A clinician-judged technique for quantifying dysarthric speech based on single-word intelligibility. <u>Journal of Communication Disorders</u>, <u>13</u>, 15-31.

Appendix A

Informed Consent

I, _____, consent to serve as a listener in this research project on measuring speech intelligibility of children.   I understand that the study involves giving an estimate of speech intelligibility in the form of a percentage and orthographically transcribing 12 speech samples. It will take a total of approximately 7 hours to listen to and transcribe the 12 samples.

I understand that participation in this study will present no physical, social, economic, or other risks. All data obtained during the course of the study will remain confidential. Published data and public records will not reveal my name.

It has been explained to me that the purpose of the study is to learn if the method of preprofessional ear estimation is an accurate and reliable measurement of speech intelligibility. I may not receive any direct benefit from taking part in this study, but my participation may help to increase knowledge which may benefit others in the future.

Kristi Mowe has offered to answer any questions I may have about the study and what I am expected to do. I understand that I am free to withdraw from participation in

this study at any time without affecting my course grade or relationship with Portland State University.

I have read and understand the above information and agree to participate in this study.

Date:_____

Signature:_____

If you have concerns or questions about this study, please contact the Chair of the Human Subjects Research Review Committee, Office of Research and Sponsored Projects, 105 Neuberger Hall, Portland State University, 503/725-3417.

Appendix B

Listening Instructions

Listener name:_____   Listener #:_____

    You will be listening to 12 connected speech samples
from children ranging in age from 4:1 to 5:1. They will be
presented in random order according to severity of
intelligibility and age.
    Please use as much objectivity as possible while
listening, however do not track or make any notes during
the speech sample. When each 3 minute sample has been
completed, you will have approximately 1 minute to evaluate
what you have heard, and select a percentage of
intelligibility between 0-100% for that speaker. Put your
estimated percentage of intelligibility in the space
provided which correlates with the speech sample presented.
    Your name will not be used in any publications or
public files related to this study. You will be referred
to by listener number only. Do you have any questions?

Sample 1 _____%

Sample 2 _____%

Sample 3 _____%

Sample 4 _____%

Sample 5 _____%

Sample 6 _____%

Sample 7 _____%

Sample 8 _____%

Sample 9 _____%

Sample 10 _____%

Sample 11 _____%

Sample 12 _____%

Appendix C

Instructions for Orthographic Transcription

Please write your name, number, speech sample and subject number you are working on in the space provided. Write down or type the words you think the child said on the lines numbered 1 to 20. Please write down one utterance per numbered line. You will not use all the lines for each child. Do not write down any interjections or filler words such as no, oh, yeah, yup, uh huh, etc.

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

Listener name:_____ Listener #:_____

Speech sample #:_____ Subject #:_____

1._____

2._____

3._____

4._____

5._____

6._____

7._____

8._____

9._____

10._____

11._____

12._____

13._____

14._____

15. _____

16. _____

17. _____

18. _____

19. _____

20. _____

Appendix D
Raw Data by Listener

Speaker #	Listener 1		Listener 2		Listener 3		Listener 4	
	E	OT	E	OT	E	OT	E	OT
	Difference		Difference		Difference		Difference	
5	98	96	100	93	98	95	99	94
	-2		-7		-3		-5	
7	95	91	95	71	90	94	99	90
	-4		-24		+4		-9	
8	98	99	98	85	98	96	98	91
	+1		-13		-2		-7	
11	78	67	70	51	65	73	65	67
	-11		-21		+8		+2	
12	83	69	55	71	75	75	75	75
	-14		+16		0		0	
16	75	54	50	37	65	45	50	40
	-21		-13		-20		-10	
17	88	79	75	63	65	71	64	73
	-9		-12		+6		+9	
19	65	48	45	32	20	43	35	52
	-17		-13		+23		+17	
20	60	46	40	33	20	33	4	24
	-14		-7		+13		+20	
21	98	98	90	87	100	90	100	90
	0		-3		-10		-10	
25	95	89	85	62	60	81	95	87
	-6		-23		+21		-8	
26	88	82	85	70	70	83	89	83
	-6		-15		+13		-6	

Note: E = Estimation Method; OT = Orthographic Transcription

Appendix E

Raw Data by Listener

Percentage of words understood in continuous speech

Estimation

	L - 1	L - 2	L - 3	L - 4	Median	Range
Speaker #						
5	98	100	98	99	98.75	2
7	95	95	90	99	94.75	9
8	98	98	98	98	98.00	0
11	78	70	65	65	69.50	13
12	83	55	75	75	72.00	28
16	75	50	65	50	60.00	25
17	88	75	65	64	73.00	24
19	65	45	20	30	41.25	45
20	60	40	20	4	31.00	56
21	98	90	100	100	97.00	10
25	95	85	60	95	83.75	35
26	88	85	70	89	83.00	19
Listener Mean	85.08	74.00	68.80	72.75		

Note: L = Listener

Appendix   F

Raw Data by Listener

Percentage-of-words understood in continuous speech

Orthographic Transcription

Speaker #	L - 1	L - 2	L - 3	L - 4	Median	Range
5	96	93	95	94	94.50	3
7	91	71	94	90	86.50	23
8	99	85	96	91	92.75	14
11	67	51	73	67	64.50	22
12	69	71	75	75	72.50	6
16	54	37	45	40	44.00	17
17	79	63	71	73	71.50	16
19	48	32	43	52	43.75	20
20	46	33	33	24	34.00	22
21	98	87	90	90	91.25	11
25	89	62	81	87	79.75	27
26	82	83	70	83	79.50	13
Listener Mean	76.50	62.90	73.25	72.16		

Note: L = Listener

Appendix G

Individual Speaker Data

Speaker #	Intell. % Est.(mean)	Intell. % OT(mean)	Difference
5	98.75	94.50	4.25
7	94.75	86.50	8.25
8	98.00	92.75	5.25
11	69.50	72.50	5.00
12	72.00	72.50	.50
16	60.00	44.00	16.00
17	73.00	71.50	2.50
19	41.25	43.75	2.50
20	31.00	34.00	3.00
21	97.00	91.25	5.75
25	83.75	79.75	4.00
26	83.00	79.50	3.50

Note: Intell. % = Intelligibility Percentage;

Est. = Estimation; OT = Orthographic Transcription

Appendix H

Raw data by Listener

Speaker (in order from most to least intelligible)	L - 1 Estim/OT		L - 2 Estim/OT		L - 3 Estim/OT		L - 4 Estim/OT	
8	98	96	98	85	98	96	98	91
5	98	96	100	93	98	95	99	94
21	98	98	90	87	100	90	100	90
25	95	89	85	62	60	81	95	87
26	88	82	85	70	70	83	89	83
7	95	91	95	71	90	94	99	90
17	88	79	75	63	65	71	64	73
12	83	69	55	71	75	75	75	75
11	78	67	70	51	65	73	65	67
16	75	54	50	37	65	45	50	40
19	65	48	45	32	20	43	35	52
20	60	46	40	33	20	33	4	24
Mean =	85.0	76.5	74.0	62.9	68.8	73.25	72.75	72.1

Note: L = Listener; Estim = Estimation Method; OT = Orthographic Transcription

Appendix I

<u>Factors that Affect Intelligibility</u>

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	Pragmatics
Articulation	Pronunciation
Communicative Disfluency	Rate
Disfluency Redundancy	Resonation
Inflection	Rhythm
Juncture	Semantics
Mean Length of Utterance	Stress
Morphology	Syntax
Morphophonemics	Voice Quality
Pauses	
Physical Posture	
Pitch	

Source: <u>Weiss Intelligibility Test</u>, Curtis E. Weiss (1982).