Portland State University

PDXScholar

Dissertations and Theses

Dissertations and Theses

1998

Speech Intelligibility of 4-Year-Old Children

M. Jane Firestone
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds

 Part of the Speech and Rhetorical Studies Commons

Let us know how access to this document benefits you.

Recommended Citation
Firestone, M. Jane, "Speech Intelligibility of 4-Year-Old Children" (1998). *Dissertations and Theses.* Paper 6322.
https://doi.org/10.15760/etd.8176

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. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

THESIS APPROVAL

The abstract and thesis of M. Jane Firestone for the Master of Science in Speech Communication: Speech and Hearing Science were presented February 25, 1998, and accepted by the thesis committee and the department.

COMMITTEE APPROVALS:

[redacted]

Mary Gordon-Brannan, Chair

[redacted]

Carla Dunn

[redacted]

Tucker Childs
Representative of the Office of Graduate Studies

DEPARTMENT APPROVAL:

[redacted]

Peter C. Ehrenhaus, Acting Chair
Department of Speech Communication

ABSTRACT

An abstract of the thesis of M. Jane Firestone for the Master of Science in Speech Communication: Speech and Hearing Science presented February 25, 1998.

Title: Speech Intelligibility of 4-Year-Old Children

While speech-language pathologists routinely measure and report speech intelligibility when assessing young children with speech impairments, normative data have not been available for comparison purposes. When assessing children to determine if their communication abilities are at or below that of peers, one must first know what the normative standards are. Knowing the normal distribution for speech intelligibility at several ages would allow for more precise uses of the intelligibility information than is currently possible. Only a few available studies exist to allow tentative normative comparisons of speech intelligibility data (Vihman & Greenlee, 1987; Ware, 1996).

The goals of this pilot study were to collect normative data for normal 4-year-olds, +/- 2 months, and to test procedural aspects of eliciting speech and determining speech intelligibility with listeners. Fifteen subjects were recruited in the greater Portland area. All were screened for normal hearing, expressive and receptive language, and phonological / articulation development. All subjects spoke English in the home and were reportedly free of motor, neurological, or developmental disorders. During the initial screening, all of the selected 15 subjects displayed

normal behavior with no significant speech-language deviations.

The investigator recorded 15 speech samples on digital audiotape in a soundproof booth at the Portland State University Speech-Language and Hearing Clinic. The 15 samples were shortened to 100+ words and compiled on separate listening tapes. Three experienced graduate clinicians in speech-language pathology listened to the tapes with varied listening order. Listeners wrote full orthographic transcriptions, using slash marks for unintelligible words.

Once all samples were transcribed, the investigator compared them after determining a starting point, and took 100 continuous words from each subject to represent a sample. A word was counted as intelligible when at least two out of three listeners transcribed it the same. Speech intelligibility percentages were derived and compared to establish normative data for 4-year-olds, +/- 2 months. One outlier percentage was removed from the data because the speech intelligibility percentage for that child was 28 percentage points lower than any other. The resulting sample size was 14. The mean intelligibility percentage was 94%, the median was 96%, and the standard deviation was 5.7.

SPEECH INTELLIGIBILITY OF 4-YEAR-OLD CHILDREN

by

M. JANE FIRESTONE

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

MASTER OF SCIENCE
in
SPEECH COMMUNICATION:
SPEECH AND HEARING SCIENCE

Portland State University
1998

ACKNOWLEDGMENTS

I want to thank Dr. Mary Gordon-Brannan for her inspiration and advise in developing this thesis and Dr. John Tetnowski for providing needed technical support. I also thank my husband and son, Gary and Elliot Firestone, and my mother, Vega Burns, for their patient listening and willingness to discuss new ideas.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION AND STATEMENT OF PURPOSE

In assessing communication disorders, speech-language pathologists (SLPs) rely on speech intelligibility measurement as a simple but useful evaluation tool. A speech intelligibility measurement is used to identify children whose speech development is substantially different from their peers. It can also be used to establish baseline measurement against which future comparisons can be made and to identify possible intervention objectives (Bernthal & Bankson, 1993; Gordon-Brannan, 1994).

Hodson and Paden (1991) defined intelligibility as the degree to which a person is understood. Intelligibility depends on factors of prosody, contextual cues, sound production, language usage, the familiarity of the listener with an individual's speech, and familiarity with the topic being discussed. Intelligibility is significant because it serves as an index to the success of expressive communication efforts. Kent (1992) stated that intelligibility is "an immediate principal criterion by which we judge a communicative attempt" (p. 1).

Speech intelligibility has long been an area of interest to SLPs when determining if a communication disorder exists in children or adults. Van Riper (1954) provided a standard for general identification of a speech disorder, stating that it must be recognizably "conspicuous, unintelligible or unpleasant" (p. 19). The relevance of intelligibility measures is not limited to early theoretical models in speech-language pathology, however. Intelligibility is a very practical means of determining the severity of disorders. Some severe speech disorders that seriously disrupt intelligibility are apraxia of speech, dysarthria, speech abnormalities secondary to cleft palate, aphasia,

laryngectomy, and severe-to-profound phonological disorders (Connolly, 1986; Kent, 1992; Schiavetti, 1992). Shriberg and Kwiatkowski (1982) found that speech intelligibility is the first factor that both SLPs and undergraduate raters considered when determining severity of speech disorders in children. Beukelman and Yorkston (1979) viewed intelligibility as an important functional index of a person's communicative skill because it integrates so many areas of competence. For these reasons, SLPs often measure or estimate speech intelligibility when conducting a full, individual speech-language assessment or general screening (Gordon-Brannan, 1994; Paul, 1995; Shriberg & Kwiatkowski, 1982).

Although intelligibility measurements have well-established value to SLPs, no developmental normative standards for speech intelligibility exist. Having clearly defined age-referenced normative data would allow objective, quantified measurement in intelligibility and might extend the current application of intelligibility measures to new areas.

Statement of Purpose

Measurements of speech intelligibility are used in the field of speech-language pathology for rating the severity of speech disorders, assessing baseline measurement before treatment, assessing progress during treatment, and evaluating the effectiveness of various interventions by noting overall intelligibility gains. The current lack of normative data for speech intelligibility limits the descriptive value of intelligibility data, leaving only impressions of relative severity without a normative benchmark. Having normative data on intelligibility would allow better global and functional characterizations of

speech, since intelligibility requires a great degree of listener integration of communicative abilities. Although normative data must eventually be collected for a wide range of ages, this pilot study will initially examine intelligibility of normal 4-year-olds as represented by a sample of 15 subjects who have typical speech and language development. The purpose of this study is to characterize the speech intelligibility of 4-year-olds. The research question is: What is the level of speech intelligibility for a typically developing 4-year-old?

Definition of Terms

Dialect: an alternate form of a language that, while differing from the most prevalent form, is not sufficiently differentiated to be considered its own separate or distinct language (Nicolosi, Harryman, & Kresheck, 1983).

Dysarthria: a group of neurologically-based motor speech disorders "manifested as disrupted oral communication caused by paralysis, weakness, abnormal tone, or incoordination" (Rosenbek & LaPointe, 1985, p. 97).

Equal-appearing scale: also known as a *Likert scale*; a method of comparison rating whereby numbers in a linear continuum are assigned to each response being compared. Typically 5-, 7-, or 9-point scales represent the full range of the possible responses (Schiavetti, 1992).

Fluency: "production of utterances in connected sequences without any extraneous pauses, hesitations, or repetitions" (Nicolosi et al., 1983, p. 99).

Intelligibility: the degree to which a person is understood, depending on factors of prosody, contextual cues, sound production, language usage, and the familiarity of the

listener with the speech or subject matter (Hodson & Paden, 1991).

Normative data: information "derived from a representative sampling of median achievement of a large group; offers a range of values against which individual comparisons can be made" (Nicolosi et al., 1983, p. 161).

Otitis media: "inflammation, usually due to viral or bacterial infection of the middle ear . . . ."(Martin, 1981, p. 296).

Perceptual measurement: a description of auditory input that takes place after the audible unit (phoneme, word, sentence, narrative) has already been received; not a measurement utilizing instrumentation (Nicolosi et al., 1983).

Phoneme: "[the] shortest arbitrary unit of sound in a given language that can be recognized as being distinct" (Nicolosi et al., pp. 181-182).

Phonological deviations: also known as *phonological processes* or *phonological errors*; systematic differences between a sound or sound sequences produced by a child and the adult standard for the sound or sound sequence. Usually sounds are modified in a process of simplification (Weiss, Gordon, & Lillywhite, 1987.)

Pragmatics: " a set of rules governing the use of language in context. . . [with] context treated as an integral part of language structure rather than a cause of language" (Nicolosi et al., 1983, p. 189).

Suprasegmental features: also known as *prosody* or *non-segmental features*; speech characteristics that include fluency, duration, rate, pitch, loudness, juncture, and rhythm; features that affect more than single vowels or consonants (Nicolosi et al., 1983).

CHAPTER II

REVIEW OF LITERATURE

All effective oral communication relies upon listener comprehension. Of the many factors contributing to speech intelligibility, some involve people and others involve the acoustic environment. First, speakers formulate messages using their knowledge of language morphology, syntax, pragmatics, semantics, and phonology. Deficiencies in the speakers' linguistic systems are usually realized in their oral expression and will often detract from intelligibility. Speech intelligibility is usually better if the speaker also observes pragmatic conventions when choosing an appropriate topic and makes conventional comments. Listeners also play an active role. To comprehend, listeners must attend to a message, have adequate hearing, and share a common language with the speaker. Further, the speaker and listener must interact in an acoustically adequate environment. A quiet physical environment may help with a message transmission while poor acoustic conditions may detract.

Effect of Unintelligible Speech

For children with unintelligible speech, their communication problem may have pervasive effects in various areas of life. Children who are significantly less intelligible in speech than their peers are considered disabled by definition (Shriberg & Kwiatkowski, 1982). Depending on age, intelligence, and other factors, children having unintelligible speech may also be handicapped. A handicap represents the degree of social, educational, or vocational disadvantage that children may experience as the result of a disability (Shriberg, 1980). Unintelligible speech is believed to have a negative impact on both psychosocial development and academic success although most research in this area

has focused on specific types of speech disorders. Shriberg and Kwiatkowski (1982) found that a significant percentage of preschoolers with low intelligibility required special education services through elementary school. Examining the broader group of children with language delays and/or speech disorders, Baker, Cantwell, and Mattison (1980) reported they have a higher incidence of attention deficits, excitability, irritability, and solitary behavior. Practicing clinicians in the schools often provide counseling support to students for issues secondary to their communication disorders (Cornett & Chabon, 1988). Any impediment to communication also affects social relationships. A lack of clear, intelligible speech beyond a certain age might result in avoidance by one's peers and frustrated adult listeners.

In the academic realm, early success in reading is partially dependent upon phonological awareness (Swank, 1994). The ability to encode intelligible speech and to decode reading both depend upon the integrity of a child's phonological system. Webster and Plante (1992) found that children with moderately to severely unintelligible speech were significantly below average in three phonological awareness skills: pseudo word reading, sentence-word segmentation, and word-phoneme segmentation. Their hypothesis is that phonological awareness and productive phonology for speech are linked in some complex manner. The exact relationship between speech intelligibility and reading achievement remains undefined, but speech intelligibility norms might be also useful in formulating screening protocols to identify children at risk for future reading difficulties.

The Developing Linguistic System

Children's speech intelligibility reflects the state of their developing linguistic systems. The basic components of language are phonology, syntax, morphology, pragmatics, and semantics. Most aspects of language will develop for years beyond the fourth birthday. Among normal children, elements of language are synthesized in a fluent and effortless manner during speech. If there is an underlying disorder, difficulties in any area may detract from speech clarity and intelligibility.

Syntax and Morphology

Children are developing and applying grammatical rules to their language throughout early childhood and elementary school age. A typical 4-year-old may have learned most of the basic syntax and morphology, but is not expected to know it all. Basic grammar is typically mastered at the level of 90% accuracy at age 5 (Owens, 1992), although 5-year-olds still must later master morphophonemic irregularities such as the variant pronunciation of the *-s* plural (Berko, 1958), comparative and superlative adjective forms (Carrow, 1985), and many derivational morphemes of increasing complexity (Wiig & Semel, 1984).

Researchers have long noted the relationship in children between the linguistic complexity of an utterance and the phonological characteristics of the utterance. Generally, a child's fluency, the ability to make well-formed utterances without interruptions or breakdowns, tends to decrease as the length and complexity of the utterances increase (Gaines, Runyan, & Meyers, 1991). In theoretical discussions regarding the speech fluency disorder of stuttering, the dynamic interaction between a

speaker's abilities (capacities) and the demands of a communication situation is known as "the demands and capacities model" (Adams, 1990, p. 136). Toddlers and preschoolers are continuously learning and integrating new linguistic information into their current linguistic systems. Extending the application of the demands and capacities model beyond fluency to other aspects of language formulation, one can predict that the speech intelligibility of children might fluctuate as their linguistic systems are evolving. Indeed, research by Vihman (1988) showed that intelligibility decreases as syntactic complexity increases in 3-year-old subjects. Panagos and Prelock (1982) found that children often use simplified syntax as the phonological complexity of an utterance increases. Conversely, an increase in linguistic proficiency often accompanies an increase in overall intelligibility (Panagos & Prelock, 1982; Paul & Shriberg, 1982).

Segmental and Suprasegmental Features

Deviations in speech may be either segmental or suprasegmental in nature. Segmental features are those parts of speech that can be broken into smaller discrete units. In the field of speech-language pathology, a number of significant studies are available that describe the development of children's speech in terms of phoneme segment acquisition. At 4 years, the phonemes of English are mostly in place but are not mastered. In many normative studies, researchers have attempted to describe the acquisition order of phoneme segments in the normal population (Poole, 1934; Prather, Hedrick, & Kern, 1975; Smit, Hand, Freilinger, Bernthal, & Bird, 1990; Templin, 1957; Wallace, 1971). Results of these studies have shown general, but not specific, agreement in the acquisition order. Several authors have stated that the phonological system is

finally perfected among typically developing children at age 6 or 7 (Ingram, 1976; Olmsted, 1971; Templin, 1957), while others have concluded the system is not in place until age 9 or later (Smit et al., 1990).

Although accuracy in phoneme production may be the most significant factor in speech intelligibility, Shriberg and Kwiatkowski (1982) discovered that other factors play significant roles, as well. They calculated percentage of consonants correct and percentage of intelligible words in continuous speech samples of 60 children with phonological disorders and found that intelligibility ratings were only moderately correlated (r=.42) to consonant accuracy. This finding is supported by related research into intelligibility of persons who are hearing impaired. Smith (1975) compared intelligibility measurements of individuals who were hard of hearing to matched subjects having the same approximate rates of segmental error. Smith found that the differences in intelligibility ratings between subjects matched for total segmental errors could vary up to 30%. The presence of non-segmental deviations explains at least part of the discrepancy.

Suprasegmental errors such as rhythm, duration of speech sounds and pauses, syllable stress, changes in pitch (frequency), intonation, and voice quality are known to contribute a great deal to listener comprehension (Weismer & Martin, 1992). Huggins' (1977) experiments with synthesized speech demonstrated that listeners can be so confused by abnormal pitch and timing that they are unable to interpret sentences in which all the sound segments are correct. Speech disorders that effect voice or volume also have the potential to disrupt intelligibility significantly. Common nondevelopmental

disorders affecting intelligibility include voice problems resulting in hoarseness and inadequate volume.

Shriberg and Kwiatkowski (1982) compared overall intelligibility, segment or phonemic accuracy, suprasegmental features, and overall speech severity ratings in the speech of children with phonological disorders. Their rating of intelligibility was derived from a ratio of intelligible words to total words spoken. The suprasegmental features examined were pitch, loudness, quality, phrasing, stress, and rate. They found that, for certain children, disorders in phrasing and rhythm were even stronger predictors of overall speech intelligibility than segmental accuracy.

The presence or absence of phonological deviations also affects the intelligibility of young children's speech. Phonological deviations are systematic oversimplications of the standard phonology of a language that result in segmental changes. Such oversimplifications are found in the children's speech before the phonological system has matured (Weiss et al., 1987). The speech of typically developing children will have some phonological deviations until about age 4, when a more standard phonology emerges (Shriberg, 1980). Hodson and Paden (1981) compared the phonological deviations of 4-year-old children with unintelligible speech to the speech of 4-year-old children who were intelligible. Although they found that phonological deviations existed in the speech of all 4-year-olds, both with and without speech disorders, the errors of the unintelligible children differed from the control group. The children who were unintelligible had a larger total of phonological deviations and more individually unique error patterns.

Both research and theoretical models often identify age 4 as a point of change

for phonological and articulatory development. Ingram (1976) accounted for the transformations occurring in normal phonological development from birth to 4+ years by a series of stages. Stage I includes prelinguistic vocalizations; stage II, the phonology needed for the first 50 lexical items; stage III, the phonology of simple morphemes that generalize from the stage II group; and stage IV, the completion of the inventory. Simple morpheme mastery is expected from 1:6 to 4:0, with the absolute completion of the phonetic inventory occurring after age 4. Ingram speculated that the addition of longer and more complex vocabulary burdens a young child's phonological system, and adversely affects intelligibility. In a child's dynamically changing linguistic system, Ingram noted that single words may not be consistently articulated. Clear articulation may be evolving, but unintelligible portions may challenge a listener until sounds and words are truly learned.

Syllable and Word Formation

Segments combine to form syllables, and syllables combine to form words. Some factors interacting and affecting speech intelligibility reside in these higher levels of the phonological and language hierarchy. Certain deviations in segment combinations at the syllable or word level reduce intelligibility more than others. Researchers in child language development have found that consonant cluster simplification (Hodson & Paden, 1981; Shriberg, Kwiatkowski, Best, Hengst, & Terselic-Weber, 1986), errors in multi-syllabic words (Shriberg et al., 1986), and errors in unstressed function words (Campbell & Shriberg, 1982) disrupt intelligibility more than other error types.

Message Formulation

Paralinguistic factors such as speech pragmatics also contribute to intelligibility. One factor in speech intelligibility is the child's mastery or observance of the conventional pragmatic functions of speech. When a speaker adheres to the usual pragmatic constraints that apply to a conversational situation, the listener has certain semantic expectations or schema to allow maximal comprehension (Grice, 1989). If a speaker speaks in an illogical or improbable manner, the chance of misunderstanding greatly increases and the intelligibility of speech is often lower. The differential between an adult's fully developed pragmatic system and the immature pragmatic abilities of children often limits an adult's comprehension and thus affects intelligibility. One example of a common pragmatic mismatch occurs when a young child switches topics erratically or frequently during a conversation. An adult listener often finds such conversations challenging to comprehend because the child has not mastered the necessary pragmatic skill of topic maintenance.

Another factor possibly affecting speech intelligibility for young children is in the semantic-pragmatic area. In new situations, children must encode information that has never been required of them before. Yoder and Davies (1990) studied the intelligibility of children having developmental delays and found that routine speech situations elicited speech that was more intelligible. Novel speech situations led to fewer intelligible utterances. These findings may or may not be applicable to the speech of children having normal phonological development.

Concomitant Disorders

Although any speech or language disorder may occur as an isolated phenomenon, high rates of concomitant speech and language disorders are reported among children (Shriberg, et al., 1986). One disorder in isolation may be mild, but in conjunction with other anomalies in language or speech, difficulties may be compounded. Having abnormal nasal resonance secondary to a cleft palate might lessen intelligibility, but having abnormal nasal resonance *and* a developmental language delay would affect intelligibility even more.

Between Listener and Speaker

Any measure of speech intelligibility is not meaningful until factors external to the speaker are also taken into account. Intelligibility is influenced by interactions between speaker and listener. These communication interactions depend on the efficiency of a speaker's transmission system, the efficiency of the listener's reception system (Schiavetti, 1992), and the underlying linguistic competence of the speaker. Although intelligible speech will normally result in listener comprehension, some intelligible statements are not effective communication due to ambiguity or contextual uncertainty (Connolly, 1986). Clearly spoken words may not always result in listener comprehension, so clear speech will not be fully comprehensible all of the time.

Some researchers have asserted that a listener's familiarity with a child speaker will improve the rate of speech intelligibility, but that the variation among individual listeners is large (Goehl & Martin, 1987). Research by Kwiatkowski and Shriberg (1992) found mothers have no advantage over trained clinicians when trying to decipher

unintelligible speech. In contrast, Flipsen (1995) found mothers to have a significant advantage over other listeners with child speakers up to age 4.

Also, listener familiarity with the message or subject matter during intelligibility measurement will inflate intelligibility estimates (Monsen, 1983; Yorkston & Beukelman, 1978, 1980). Besides the familiarity effect, the experience of the listener may affect intelligibility ratings. Experienced listeners, those listeners who have had exposure to the speech of many children, may rate speech intelligibility higher than inexperienced listeners would. The difference between inexperienced and experienced raters often disappears when listeners are asked to rank speakers in order, however. When inexperienced and experienced raters made an ordinal ranking of subjects' speech in direct magnitude estimation, researchers noted no difference between the groups (Schiavetti, Metz, & Sitler, 1981).

Social and environmental factors may also play a role in rates of speech intelligibility. In American English, differences are noted in major dialect groups such as Black English, southern white nonstandard English, and Appalachian English. Wolfram and Fasold (1974) noted that although nonstandard dialects share most features in common with standard American English, certain features in the lexicon and syntax might differ. These dialects exist in a "continuum of divergency" (Wolfram & Fasold, p.33). If a speaker uses one or more divergent features in communication, the listener belonging to another dialect group may not fully comprehend the message. A message that is not fully comprehended is, by definition, not fully intelligible.

The quality of the physical environment also plays a role. Having the opportunity

to see a speaker's face may improve intelligibility. The presence of ambient noise or other intruding conversation in the speaking environment may diminish it.

Listener and Context

Listeners can comprehend speech in part because of their ability to use common sense and inference to interpret what is said. Predictability and redundancy are two important factors that allow successful inference to take place. Predictability allows even partially garbled messages to be correctly identified, e.g., *He hit the ball with a b. . . . (bat).* Redundant information in a message also allows a listener to infer correctly despite errors, using other information that is present, e.g., *They lost all three dog. . . . (dogs).*

Although elements within a message such as redundancy and predictability may interact to enhance comprehension, confusing contextual elements may detract from it. Errors create more intelligibility problems because a listener relies on what was already said when interpreting new information. Weston and Shriberg (1992) suggested that unintelligible segments directly adjacent to a questionable word may decrease the likelihood of the questionable word being correctly identified.

Normative Studies

Relatively few normative speech intelligibility studies are available for comparison. Vihman and Greenlee (1987) observed that 10 normally developing 3-year-olds had a wide range of speech intelligibility. Reseachers used two methods of intelligibility assessment: (a) having 3 coders calculate the proportion of unintelligible utterances to the total number of utterances, and (b) rank ordering children based on

composite scores. The results were that all 10 children were greater than 50% intelligible, with a range of intelligibility from 54% to 80%. The 3-year-olds using complex syntax were the hardest to understand.

Assessing 4 and 5-year-olds with normal phonological deviations, Gordon-Brannan (1993) reported a 93% mean for speech intelligibility with a range of 73% to 100% using the percentage of intelligible words as the standard for measurement. Other sources have concluded that 4-year-olds are generally intelligible; however, they have not cited specific data to support that conclusion (Grunwell, 1992; Weiss & Lilywhite, 1981). Weiss and Lillywhite (1981) proposed intelligibility norms for young children from 0 to 48 months (Table 1).

Table 1

Intelligibility Norms Reported by Weiss and Lillywhite (1981): 0 to 48 Months

Age in Months	Speech Intelligibility
0-18	0-25%
19-24	26-50%
25-30	51-70%
31-36	71-80%
37-42	81-90%
43-47	91-99%
48	100%

Although this information suggests empirically derived norms, the authors offered no explanation of the source. These may be estimates or guidelines derived from clinical observation or related research. The accuracy of these standards is yet to be demonstrated. Especially notable in Weiss and Lillywhite's (1981) normal intelligibility ratings is the lack of an anticipated range in speech intelligibility for four-year-olds. Vihman and Greenlee (1987) reported a large range (54-80%) in the speech intelligibility of normal 3-year-olds, but Weiss and Lillywhite (1981) anticipated 100% intelligibility by 4 years.

Measurement Issues

SLPs often assess overall speech intelligibility in medical and school settings. Intelligibility measures have been especially important in specific specialized assessments of persons who have a hearing impairment, foreign accent, dysarthria, or a developmental speech disorder. While the interpretation and application of intelligibility data may differ among these disparate areas, certain common points arise in the discussion of intelligibility measurement for research and intervention.

A rating of speech intelligibility determination is a perceptual measurement, usually made by a clinician-listener (Yorkston, Beukelman, & Bell, 1988). The value of a perceptually-based intelligibility rating lies in the human ability to integrate the perceptual data with meaning. A successful outcome is speech comprehension. Speech intelligibility may be subjectively or objectively determined. Both subjective and objective measurements have value, depending on the needs of the clinician (Yorkston & Beukelman, 1978).

Subjective Measurement

Gross estimated percentages of intelligibility and equal-appearing rating scales are both common subjective measures. Estimated percentages of intelligibility have value when communicating with parents or allied professionals because of the speed of assessment and communicative simplicity (Yorkston & Beukelman, 1981). Equal-appearing scales require a listener to make another global assessment concerning the quality or comprehensibility of a speaker's utterances. After a sample is heard, the listener chooses an overall numerical value with a descriptor from a rating scale. The National Technical Institute for the Deaf (NTID) developed one such scale. The listener assigns a value of *5* to the speech of a child who is fully intelligible and *1* for one who is totally unintelligible, with values of *2* to *4* lying between (Johnson, 1975.) A major shortcoming has been found in the NTID scale, however. Two studies found that speakers with middle range ratings (2 - 4) on the NTID had unacceptably high standard deviations from the mean when NTID results were compared to objective ratings from word identification tasks (Kelly, Dancer, & Bradley, 1986; Samar & Metz, 1988). These findings suggest that most listeners can make gross perceptual ratings of intelligibility at only a few levels (Schiavetti, 1992).

Objective Measurement.

Two approaches to objective speech intelligibility rating are common. The first approach involves formalized repetition or completion tasks. In repetition, words and sentences are repeated after the examiner. In a completion task, the examiner asks subjects to complete phrases or words to form a sentence. Another second objective

approach involves calculating the proportion of unintelligible words to total words in connected speech, without imitating the examiner. When compared to estimates of speech intelligibility, objective measurement techniques are superior in test-retest reliability. Both objective tests and estimations of intelligibility rank order a group of subjects similarly, however (Yorkston & Beukelman, 1978).

In objective measurement, word and sentence repetition tests are often favored for ease of administration. Such tests use a closed-set method: eliciting answers from a closed-set of words or possible responses. The Preschool Speech Intelligibility Measure (PSIM) is one such test recently designed for use with young children (Morris, Wilcox, & Schooling, 1995). The PSIM was based on a single-word intelligibility test devised by Yorkson and Beukelman (1980, 1981) for dysarthric adults. In the PSIM, the examiner models and the child repeats the words while being tape-recorded. Other examiners listen and evaluate the recordings later. While initial data from the PSIM development sound promising, the protocol requires a minimum of two examiners: one to present words and another to listen. Additionally the PSIM has less face validity than the calculation of intelligible words because the client is receiving extensive cueing.

Speech samples taken from connected speech are commonly used to make intelligibility determinations. Many researchers and theorists find that connected speech provides face validity to intelligibility measurement that is lacking in simple naming or repetition tasks (Gordon-Brannan, 1994; Morrison & Shriberg, 1992; Weston & Shriberg, 1992). Moreover, an intelligibility measurement derived from connected speech preserves communication functionality and thus provides more naturalistic data.

Connected speech demonstrates the ability of a speaker to integrate multiple speech, language, and paralinguistic capabilities.

Comparatively, SLPs rely less on speech intelligibility assessment than on standardized tests that examine several areas of linguistic competence. Those areas most widely examined are phoneme production (articulation), vocabulary, and other selected skills in language such as listening comprehension and syntax. Speech intelligibility assessment from connected speech has some advantages over most of the aforementioned tests because intelligibility measures sample integrated skills rather than isolated ones. Second, a speech sample allowing spontaneous comments is relatively more natural than responses derived from a closed set question and answer paradigm.

Summary and Conclusion

Speech intelligibility is a multifaceted area of speech and language competence, important because it integrates communicative skills across many domains: linguistic, paralinguistic, social, and contextual. Although speech intelligibility measurement has been routinely performed by SLPs during evaluation of children and adults, normative standards have not been generally available. Lacking empirical data as a basis for comparisons, intelligibility measurement has been largely impressionistic, allowing for only broad interpretations of possible severity or age appropriateness. Only a few studies to date have provided normative data by age. This pilot study contributes to the known normative data on normal speech intelligibility for 4-year-olds. This study also examined the methodological problems in speech elicitation and listener rating systems.

CHAPTER III

METHODS

The purpose of this study is to describe the normal level of speech intelligibility for children of 4 years of age, +/- 2 months, who have typically developing speech and language skills. This pilot study contributes to a future normative speech intelligibility project to be conducted by Gordon-Brannan that will include a wider range of ages.

Participants

Subject Recruitment

Parents of prospective subjects were recruited through information flyers distributed by the participating area preschools, ads in *Portland Parent* newspaper (Appendix A), and personal communication. One of the first mothers who answered an advertisement in *Portland Parent* subsequently referred 4 more qualified parents and children.

The sample for this study consisted of 15 children, 7 males and 8 females. All had met the basic selection criteria:

1. Children were at least 3 years, 10 months and no older than 4 years, 2 months.

2. All children demonstrated hearing normal hearing when tested with an audiometer at 20 dB for the frequencies of 1000, 2000, and 4000 Hz (ASHA, 1993).

3. The parent or guardian of each subject signed an informed consent agreement (Appendix B) and completed a questionnaire (Appendix C). Each parent received a general information letter describing the purpose and procedures involved in the study (Appendix D).

4. All were native speakers of English. Native speakers of English were defined

as children who speak English with one or more parents in the home. These data were reported by parents in the questionnaire (Appendix C).

5. None had known disabilities in motor, neurological, or physical areas that would affect speech or language. These data were available from responses on the parental questionnaire (Appendix C).

6. Children selected for participation in this study displayed receptive and expressive language abilities within normal limits for their age, as determined by the Fluharty Preschool Speech and Language Screening Test (Fluharty, 1978).

7. Subjects had normal phonological development as measured by The Phonological Processes - Revised: Preschool Screening (APP-R: Preschool). The APP-R: Preschool was developed by Hodson (1986) to identify children who show below normal preschool-age phonological development.

8. Children selected as subjects were observed to have normal voice and fluency abilities for their age. This was determined by clinical observation during the first screening session.

The investigator used the published normal cutoff scores for both the Fluharty and APPR: Preschool as criteria for acceptance. All subjects scored at or above the published normal cutoff levels. Two of the screening instruments measured phoneme production: the Fluharty articulation section and the APP-R: Preschool. The articulation section of the Fluharty rates articulation based on the total number of targeted phoneme errors. The APP-R: Preschool rates phonological development by noting recurrent phoneme error patterns

Twenty-one children were screened. Four children failed the initial screening: three girls and one boy. One girl failed both the Fluharty articulation section and the APP-R: Preschool. The second girl failed only the Fluharty articulation section, but passed the APP-R: Preschool. The third girl failed the hearing screening. The only boy who failed screening did not pass the comprehension section of the Fluharty and had difficulty staying on task throughout. Two other children passed their initial screenings, but their parents decided to withdraw from the study. The informational letter to parents (Appendix D) had stated that participants had the right to withdraw anytime without explanation.

The mean age of the participants was 4:0, and the median age was also 4:0. The age range of the subjects was 3:10 to 4:2. The oldest subject who was 4:2 was recorded on the exact day he turned 4 years, 2 months.

Listeners

A group of 3 graduate students majoring in Speech-Language Pathology at Portland State University were volunteer listeners in this study. The 3 listeners were compensated for their work, given the choice of dinner at a restaurant or being paid at $10 an hour. Each listener had normal hearing and had completed course work in phonetics, phonological disorders and child language development. Each listener had to have completed one or more terms as a student clinician. The requirement for each listener was to write an orthographic transcription of each of the 15 speech samples.

Instrumentation

A portable Beltone 120 pure tone audiometer with TDH 39 headphones was

used to conduct the audiometric hearing screening of each potential subject and listener. The audiological screening procedures complied with American Speech-Language-Hearing Association (ASHA) guidelines (AHSA, 1993). Speech samples were recorded at the Portland State University Speech and Hearing Clinic on a PCM-2300 Sony DAT recorder. Listeners played back the recorded digital speech samples using the same Sony digital recorder with a Phillips FA950 stereo amplifier and JBL speakers in a sound field. The output levels were adjusted by the individual listeners.

Procedures

Preliminary Screening

All of the 4-year-old subjects lived in the greater Portland area. Parents of all participants signed an informed consent agreement (Appendix B) and provided information concerning the children's speech, language, and hearing history in a brief questionnaire (Appendix C). Included were questions concerning the parents' current occupation, highest levels of education, the child's speech and language history, diagnosed problems, and history of ear infections. The answers to the questionnaire provided two kinds of information. First, certain children whose parents reported communication disorders would not have been included in the study. In fact, no child whose parent completed a questionnaire reported a disorder or history of language delay that disqualified them. A few parents reported chronic otitis media, but none were disqualified for that reason. Through an initial telephone conversation, one parent/child was prescreened and determined ineligible. That child reportedly had a repaired cleft palate and was not considered a normal 4-year-old. The data from the questionnaire

were used less as a screening instrument and more as a source of information against which the intelligibility ranking could be compared.

Children were first screened for normal hearing. Normal hearing was defined as having passed bilateral hearing screening at 20 dB HL for the frequencies of 500, 1000, 2000, and 4000 Hz (ASHA, 1993). The ambient noise levels due to traffic and appliance noise were generally too high at the preschools or homes where the testing took place to allow screening at 500 Hz.

Those who passed the hearing screening were then given specific speech and language screening tests. The Fluharty Preschool Speech and Language Screening Test is a standardized screening instrument for articulation, comprehension, and expressive language. This test has high inter- and intra-tester reliability and high content validity. Section A of the Fluharty requires the child to name common objects. Section A measures both expressive vocabulary and the accurate production of designated phonemes. In section B, a child responds nonverbally to statements or questions by the examiner. In section C, a child repeats sentences read by the examiner. Those subjects who met or exceeded the normal cutoff scores in each of the four areas (identification, articulation, comprehension, and repetition) met the language screening criteria.

The APP-R: Preschool was also administered individually to the subjects at home or at preschool as part of the eligibility screening. The test is administered by eliciting the names of common objects. Although the APP-R: Preschool is not standardized, the test is used to distinguish between children who are developing normally and those who have deviant phonology. The standard normal cutoff for the APP-R is based on the

number of errors occurring within six identified areas of phonological development: (a) consonant omissions, (b) consonant sequence reductions, (c) substitutions for stridents, (d) substitutions for velars, (e) substitutions for liquids, and (f) backing of anterior consonants. Children having speech difficulty in two or more of the six identified areas were not included as subjects in this study. Beyond its use as a screening instrument, the researcher used the APP-R: Preschool results as a short profile of phonological development for each subject. Each child completing all or part of the screening was given a small prize for participating.

Language Sample Collection

Qualified subjects were invited with their parents to PSU so that a narrative speech sample could be recorded. After a child had clearly qualified as a subject after screening, the examiner verbally invited the child to visit the "school for big kids." The examiner explained that the child could visit the school to talk about two picture books. The examiner further explained that a good tape recording could be made there in a special room. With parental approval, the examiner stated that another little prize would be offered at the end of the taping at PSU.

The investigator elicited speech samples from each subject in an acoustically controlled recording area at the Portland State University Speech-Language and Hearing Clinic. Each child was provided a few minutes of preliminary play to become accustomed to the recording booth. The investigator offered a stuffed bunny for holding and the recording booth was decorated with posters to create an inviting atmosphere. The investigator repeated the information explained earlier: the child was asked to look

at two picture books and tell (not read) the stories so that a good recording could be made of the child's speech. The picture story books, *The Relatives Came* (Rylant & Gammell, 1985) and *Good Dog, Carl* (Day, 1985), were introduced as stimulus materials. The speech sample of approximately 200 words was recorded from the story book conversation. The investigator used a restricted number of statements to elicit comments about the stories: *what is next, what is happening, tell me what this is about, what about that, now what, and then what* and brief interjections of interest, agreement, or surprise such as *wow, look (there), see (that), uh-huh, yes,* and *hmm*. The investigator also pointed silently at areas of interest on the pages if it appeared that the child had not noticed. Restricting the eliciting verbal stimuli was an attempt to avoid inadvertently providing verbal models for the subjects to copy. Verbal cues and questions from the examiner also would give clues to the listeners later.

Preparation of Listener Tapes

The researcher's goal was to provide the listeners language samples in a known context for each subject. After the subjects were taped, the 15 original speech samples were edited for length. Each sample was shortened to approximately 120-150 words, beginning with the child's tenth sentence. A 10- sentence lead was not possible for subject 4 because she quickly tired of her speaking task and that resulted in a shorter sample. The 15 shortened samples were then transferred to 3 tapes. Tape A had speech samples from subjects 1-5, Tape B had samples from subjects 6-10, and Tape C had samples from subjects 11-15.

Listening Order

The investigator assigned a different listening order for the tapes to each listener: either A-B-C, B-C-A, or C-A-B to counteract the order effect. The order effect can potentially invalidate the comparison of the listeners' performances. A listener may experience fatigue with a long task, resulting in poorer performance toward the end. Conversely, a listener's cumulative experience could improve comprehension, making the last of the samples easier to comprehend than the first. To minimize the order effect, the listening order was varied and the listening task was broken into sections. The individual language samples were identical for each listener; only the order of samples changed.

Listener Orientation

The investigator verbally reviewed the listener protocol (Appendix E) with each listener and provided the same protocol in writing. The investigator asked the listeners to review the two books used as picture stimuli before listening and to refer to them as needed in the lab while listening and transcribing. The purpose of previewing the material was to simulate listening with a meaningful context. Comparatively, a decontextualized listening task (i.e., not knowing what to expect from the speaker) is much harder and would likely yield lower ratings of intelligibility. The listening occurred in three separate days, with one listening day per tape. Each listener wrote each word using standard orthography. The listeners were permitted to turn off the recorder as needed to allow the time for writing. Listeners were allowed to review the tapes three times. The method used to accomplish three listenings was left to the listener's discretion: segments could

be reviewed as needed (for a maximum of three listenings) or the listener could listen to a whole sample or tape three times. No advantage was anticipated with one method of review over another. The items a listener had not yet understood after the third listening were marked unintelligible with a slash mark. Listeners used slash marks for unintelligible words, but wrote their "best guesses" for partially intelligible words. This procedure is justified because normal listening involves making guesses about what has been said.

Scoring and Data Analysis

The three transcriptions for each of the 15 subjects were analyzed. The researcher compared the 3 listeners' transcriptions for each sample by placing them word-by-word in a matrix so that precise comparisons were possible.. A word was deemed intelligible if at least 2 out of 3 listeners had transcribed it the same. If a word was identified by none or only one of three listeners, it was judged unintelligible. The researcher first found a word in agreement that could be a start word in all transcripts. Using the starting word as the first item in the sample, the start word and the next 100+ words from the transcripts were written in the matrix. This allowed for easy visual comparisons and word counts. Appendix I illustrates how a line describing *Good Dog, Carl* was transcribed and scored. Differences in selected function words and affixes were considered insignificant for the purpose of rating intelligibility. Transcriptions that differed only in number (e.g., *dog, dogs*), or the modifying articles (e.g., *a, an, the*) were counted as words in agreement and scored as one. Words in the same position that had the same derivational morpheme (e.g., *caring, cared*) were also counted as words in

agreement. By comparing transcripts and matching them word-by-word, the researcher noted that words and phrases appeared occasionally in one transcript but were not part of the other two. These extra words were not counted in the total word score. Certain utterances were excluded from the data: false starts, *yes / no* responses, interjections, and filler words (e.g., *hey, uh, hmm, huh*). Repeated words were included in the word count. Full agreement of all 3 listeners was not required because of the possible variability in listener experience or aptitude.

After comparing the three separate transcriptions, a continuous string of 100 words was taken from the common start point. By subtracting the unintelligible words in the sample from 100, the percentage of speech intelligibility was derived for each. These data were analyzed to determine the sample mean, median, standard deviation, and sample range for speech intelligibility for normal 4-year-old speakers.

CHAPTER IV

RESULTS AND DISCUSSION

Intelligibility

The purpose of this study was to collect normative data on speech intelligibility for 4-year-olds. Fifteen typically-developing children participated in the study as subjects. The children were tape-recorded while they described the two stories presented in picture books. During the taping, the researcher elicited the samples using as few direct verbal prompts as possible. Three listeners, already familiar with books that provided the context of the speech samples, orthographically transcribed the taped 100 word samples. The researcher calculated the percentage of intelligible speech in context by comparing the 3 transcripts and counting words as intelligible if at least 2 out of 3 listeners agreed.

Individual results for the 15 subjects are displayed in Appendix F. The descriptive statistics for speech intelligibility of the typically developing 4-year-olds in this study are shown in Table 2. The mean percentage of intelligibility for the 15 4-year-olds was 91%

Table 2

Speech Intelligibility for 4-Year-Olds: n=15

Ages	n	Mean	Range	Median	SD
3:10-4:2	15	91%	53%-100%	96%	11.9

and the median was 96%. The intelligibility scores ranged from 53% to 100% with a standard deviation of 11.9. Scores were clustered around 96%-97%: two subjects were 97% intelligible and four were 96% intelligible. Scores on the higher end of the range are spread evenly from 100% to 94%. Below the 94% level, however, intelligibility scores decrease to 86%, 85%, and 81%. The one outlier is an isolated low score of 53%.

One outlier percentage of intelligibility was 28 percentage points below that of the next lowest performing subject. This lowest measure of intelligibility was 53%, which was the percentage for subject 3, a female of 4 years. If the outlier is excluded from the calculation of the descriptive statistics, the mean is 94%, the median is 96%, and the standard deviation is 5.7 (Table 3). The range of speech intelligibility for the resulting sample of 14 4-year-olds is 81 to 100%.

Table 3

Speech Intelligibility for 4-Year-Olds Excluding Outlier: n=14

Ages	n	Mean	Range	Median	SD
3:10-4:2	14	94%	81%-100%	96%	5.7

The mean rate of intelligibility for the 7 girls is 94% and the median is 96 %. The range is 81%-99%, with a standard deviation of 6.0. The mean rate of intelligibility for the 7 boys is 94% and the median is 96%. The range is 85% to 100%, with a standard deviation of 5.9.

Table 4

Intelligibility Distribution for 4-Year-Old Girls (n=7) and Boys (n=7)

	Age	n	Mean	Median	SD	Range
Girls	3:10 - 4:0	7	94%	96%	6.0	81-99%
Boys	3:10 -4:2	7	94%	96%	5.9	85-100%

In this small sample, no difference is apparent between girls and boys at the age of 4 in speech intelligibility.

Sample Characteristics

Education. The children who participated in the study primarily came from middle class backgrounds as inferred from parental reports of their educational background on the questionnaire (Table 6). The reported highest parent educational level ranged from a high school diploma to postbaccalaureate graduate degree. All but subject 4 had at least one parent who was a college graduate. Subject 5 had a single adoptive parent, so the education of the father was not reported. The parent of subject 9 did not report the father's education but stated that he worked in business.

Table 5

Reported Highest Education of Parents

Highest Education Received	Number of Mothers	Number of Fathers
High School (H.S.) only or H.S. plus courses	1	6
B.S. or B.A.	10	6
Graduate Degree	4	1
Not Reported	0	2

Geographic and Social Factors. The parents responding to advertisements for subjects lived in Portland or the outlying cities of Tigard, Tualatin, Hillsboro, Clackamas, and Beaverton. One parent did not report an address other than that of her business. Fourteen out of 15 responding parents were mothers. Consequently, 14 out of 15 consent statements and questionnaires were completed by the mother only. Subjects 5 and 13 were children who were adopted during infancy from Vietnam and Korea, respectively. Although no income data were reported in the questionnaire, the educational achievement of the parents suggests that many children benefit from opportunities and resources not always found in the general population. One 4-year-old child in this study, for example, had his own personal computer and printer in his bedroom. This may be typical for 4-year-olds in the future, but probably represents an advantaged home environment at this time.

Health and Developmental History. Appendix F includes data from the parent questionnaire regarding ear infections and speech and language milestones. The reported number of ear infections ranged from 0 (reported by 4) to over 15 (reported by one). One parent reported that ear infections occurred "often" but did not state a number. The most frequently reported number of infections was at the 3- 4 level (reported 6 times). In the questionnaire, parents reported the age when the first word was spoken and the age of the first word combinations. Both skills are considered significant milestones in language development. Two parents reported older ages in these two areas.

Parents reported in the questionnaires that all child subjects spoke English in the home. All parents reported their children's speech was comprehensible to their own family members and to others. The developmental history of the 15 original subjects varied widely, however (Appendix G). All data were stated in months to allow for comparison. The earliest report of a word being spoken was 5 months. The latest first word reported was at 2 years. The earliest reported word combination was at 9 months. The latest was 24 + months. Subject 11 was adopted and was exposed to Korean only until 8 ½ months. Her mother did not recall when her first English word and word combinations emerged.

Discussion

Intelligibility measurements are commonly included in clinical descriptions of speech and language competence. Although intelligibility percentages are widely used, the implications given percentage rating may be unknown because of the lack of normative data. One might infer that 4-year-olds whose speech is only 50% intelligible

are below average in communication skills, but until now, there has been no empirical evidence to support that conclusion. One problem is that a SLP may calculate and report a rate of speech intelligibility (80%, for example) that could be seriously misinterpreted. The SLP may intend to communicate with this percentage that a moderate degree of impairment exists. The recipient of this information might interpret an 80% intelligibility rating as signifying a severe communication limitation. Numbers may seem precise, but lacking a comparative standard, they are truly vague. Further, SLPs may use a percentage of intelligibility and attach some verbal modifiers such as *mild, moderate, severe*, or *profound impairment* without having a factual basis for such interpretations. In the extremes, such as 100% intelligibility at age 3 or 10% intelligibility at age 5, interpretation may not be difficult. Commonly, however, intelligibility scores often fall in the middle ground where the distinctions are less clear. Another troublesome aspect of speech intelligibility ratings is that a certain amount of time and effort is necessary to obtain reliable estimates or measurement of intelligible speech, but once the data are obtained, the application of the information is very constrained. Intelligibility data provide descriptive support in overall assessment, but is rarely a decisive piece of information in an adult or pediatric speech-language profile. When more is known about the normal distribution of speech intelligibility, the data may be a more powerful descriptor in the assessment process.

The data contained in this study should provide a tentative standard for speech intelligibility in normal 4-year-olds. These normative standards allow a reported percentage of speech intelligibility to be classified as normal, abnormal, or borderline.

Once even more data is collected for 3-year-olds, 4-year-olds, 5-year-olds, and the half-years between, a reliable set of descriptive statistics will be available, allowing SLPs to make precise inferential statements concerning children's current communication abilities and to estimate growth over time. All this can be accomplished without relying on commercially prepared testing materials. All the same, the examiner must still exercise good professional judgment while eliciting a speech sample in choosing appropriate stimulus materials and making minimal verbal prompts.

Performance and Individual History

Articulation Accuracy. Speech intelligibility has often been assumed to be primarily related to accurate phoneme articulation, but when closely examined, the foundations of intelligible speech are more complex. Shriberg and Kwiatkowski (1982) showed that intelligibility was only moderately correlated to consonant accuracy in speech samples of children with phonological disorders. In these speech samples from normal 4-year-olds, the number of phoneme errors on the Fluharty and APP-R: Preschool also did not predict the ranking of children in overall speech intelligibility (Appendix F). The lowest rate of intelligibility (53%), for example, was measured in a female subject who made 4 phoneme errors on the Fluharty and no phonological deviations as measured by the APP-R: Preschool. Conversely, the male with a high rate of overall intelligibility (98%) had the largest number of phoneme errors (9) on the Fluharty and one phonological deviation as measured by the APP-R: Preschool.

Developmental History. In the parental questionnaire, several questions related to the child's early language development, asking parents to recall when their children

spoke their first word and word combinations. The accuracy of these estimates is suspect because the age range was very large for a group of normally developing children. Parents of subjects 6 and 12 reported that the first word was spoken at 24 months. Since first words generally emerge between 12 and 18 months, a first word at 24 months suggests an expressive language delay (Owens, 1992). At the other end of the spectrum, Subject 4 was reported to have spoken her first word at 5 months. Since the first reported word was *ma ma*, it is likely that this first word was not a true word. A typically developing infant will vocalize and imitate sounds in the environment from 3 months on. To qualify as a word, the sounds must resemble the adult standard and there must be a particular referent (an object or person) that is consistently implied. It is not sufficient that a child repeat sound combinations that sound like words. The sounds must have a specific referent and be intentional, rather than incidental. Typically, caregivers respond to sounds made by an infant as if the sounds were meaningful conversation (Owens, 1992). For these reasons, caregivers probably do not distinguish between true words and babbling. Brown (1973) stated that the skill of combining words usually emerges at 18 months, but parents of subjects 5, 8, and 14 reported word combinations as late as 24 months (Appendix F).

Ear Infections. In the parent questionnaires, some parents reported chronic ear infections but none reported such severe infections as to require placement of ventilation tubes. The reported number of ear infections appeared to have little relationship to intelligibility ranking (Appendix F). Parents also expressed concern over their inability to recall the exact details about ear infections.

Narrative Ability. An advantage of using an intelligibility measurement is its power to examine the integrated communication skills of the speaker by a speech sample. The timing of the measurement, however, is crucial. Such a measurement is neither valid nor reliable if a subject is nervous, uncooperative, or fearful. Although subject 3, the child having the outlier intelligibility score, appeared fairly comfortable during taping, inhibition may have contributed to her lower intelligibility rating. If one assumes that subject 3 was speaking as she habitually does, a review of her recorded sample reveals an immature narrative ability. Narrative cohesion is needed to tell a story that makes sense. Applebee (1978) has described the evolution of children's narrative abilities as beginning with heap stories. In a heap story, normal children in the age range of 2-3 years label many events with no unifying theme or organization. The syntax is usually simple with no defined temporal or causal relationships. A portion of subject 3's story book description resembles a heap story. (Slash marks represent unintelligible words.)

> */ look at them and then at the flowers / / the baby's on it. . . he opened that on the floor / messed up. He got all over it and then it spilled . . . brushing the baby and putting him back to bed . . . bedtime. and this and that and that and the mama didn't . . . and the baby and the dog. It has a new friend.*

A heap story is especially hard for an adult listener to follow without explicit contextual clues. Lacking explicit clues, the listeners could not comprehend her story very well. A large portion of her narrative consisted of naming things or pointing. Additionally, subject 3 had fluctuating volume, breathy voice quality, and singsong intonation. None of these limitations were apparent to the investigator during informal conversations on an earlier day, especially issues with volume and voice quality. The

researcher did not ask the parent's opinion about whether the child's speech was typical that day. Since this child was the third to be taped, the researcher had not formed an opinion about what to expect from children in the 3:10 - 4:2 range. In retrospect, subject 3 was comparatively shy, but cooperative. Several factors undoubtedly contributed to her lack of speech clarity during taping, but none were foreseeable.

For another subject, narrative ability may have improved speech intelligibility. Subject 6 has progressed beyond heap stories and sequence stories to the primitive narrative stage. This is the narrative style that is developmentally appropriate for this age. The primitive narrative is used by typically developing children between 4-4 ½ years. The primitive narrative was described by Applebee (1978) as a narrative built around a person or persons with some initiating event, followed by action, followed by some final consequence. No real character development is evident, but there is cohesion around a central theme. An example from subject 6 illustrates:

> *The family is going on a vacation. They driving the little car. They knocked down the garbage can. They driving down and going up a mountain. They knocked the fence down. They tangle the swing up. They're sad now. They're all in the house. They're having dinner.*

Syntactically, subject 6 had very few fragmented syntactic elements in his speech sample. Most sentences contained a subject and predicate. Although this subject makes obvious articulation errors, he is fairly easy to understand. For example, subject 6 routinely substitutes */d/* for */g/* and uses the glide /j/ for the phoneme /l/, transforming the word *gum* to *dum* and *leaf* to *yeaf*. His errors are systematic, so it is possible that listeners can recognize his speech patterns and decode words with errors.

Voice Quality. Voice quality concerns were evident from the first taping because several 4-year-old speakers enjoy using "funny voices" when telling a humorous story. A "funny voice" is a very subjective description. The impression it imparts is that the child-speaker is very involved in the story, expressing mock dismay, horror, and amusement at the appropriate times. The subjects using funny voices to narrate were subject 1, female; subject 8, female; and subject 10, male. Subject 1 was 99% intelligible, subject 8 was 86% intelligible, and subject 10 was 96% intelligible. Despite the investigator's initial concern that the funny voices would complicate the intelligibility judgments, this voice quality did not appear to tax the listeners' abilities. Other unusual voice qualities were observed. Subject 9, a male, had a vocal glottal fry along with a tendency to mumble. Glottal fry is a low voice register, produced with low air flow coming through the vocal folds in irregular bursts. The vibration of the vocal folds in glottal fry is slow and the sound is similar to popping or creaking (Colton & Casper, 1996). When speaking with subject 9 in person, the investigator found him very intelligible, although his voice quality was distinctive. The listeners transcribing the tape, however, found him 85% intelligible. He was ranked 13th out of 14 subjects. The question remains whether subject 9 should have been excluded from this sample because of his distinctive and possibly abnormal voice. In this case, his voice was not sufficiently unusual during the screening and interview to warrant his disqualification.

Methodological Concerns

Calculating Rates of Intelligibility. This study relied on agreement of 2 out of 3 listeners to determine general speech intelligibility. Taken from another point of view,

one might think that absolute unanimity would be needed to demonstrate speech intelligibility. While that might seem theoretically appealing, in the analysis of the transcripts and tapes, the investigator saw that all listeners/ transcribers failed to transcribe whole utterances at times. These omissions did not necessarily arise in the same sections. Most of the parts that were omitted were words that had been spoken rapidly. Several omissions were false starts or asides that did not pertain to the picture books. Perhaps the listeners felt that certain comments were unneeded, inappropriate, and should be left out. Another possibility is that they did not catch them. If subsequent research uses listeners to determine intelligibility, clearer instructions should be made regarding the types of utterances desired for study. Rather than seeking some samples specifically related to the pictures, here all subjects' comments were supposed to have been included. This fact should have been explicitly stated in the listener protocol.

A comparison was made between two methods of making speech intelligibility determinations using listener transcripts. In the first method (used in this study), the researcher counted a response as intelligible every time two or more listeners transcribed a words in a 100-word sample the same. Limitations applied to counting article differences and words with the same roots. A second method of determining intelligibility required agreement of all three speakers (unanimous agreement) on every word in the same 100-word sample, with the same limitations applying to articles and words with the same roots. A comparison was made between the intelligibility ratings derived in this study (requiring at minimum, agreement from two out of three listeners) with ratings derived from unanimous agreement of listeners (Appendix H). The two

methods yielded very different intelligibility ratings. For each subject, the method requiring unanimous agreement naturally yielded lower intelligibility scores. The mean intelligibility score for ratings derived from unanimous agreement was 79% compared to the 91% mean derived by two out of three in agreement. The median score derived from unanimous agreement was 84%, the standard deviation was 15.8, and the range was 33%-94%. The median score for two out of three in agreement was 96%, the standard deviation was 11.9, and the range was 53% -100%. Two intelligibility ratings were a full 20 percentage points lower using the method requiring unanimous agreement. The average intelligibility rating score was lowered by 12.5 percentage points. The result of the unanimous agreement method was that every listener's lapse in concentration affected the rating. The method using two out of three listeners resulted in a corrective balance: if one listener lost track of what was said, often the other two listeners would catch it. This outcome was more desirable because the ratings were less dependent on the fluctuating abilities of the listeners.

Table 6

Descriptive Statistics for Intelligibility Measurements: 2 of 3 Listeners and 3 of 3 Listeners

	Sample Size	Mean	Median	Standard Deviation	Range
Agreement: 2/3	*15*	*91%*	*96%*	*11.9*	*53%-100%*
Agreement: 3/3	*15*	*79%*	*84%*	*15.8*	*33%-94%*

Comparison to Other Normative Studies

Although this pilot study had only a small sample size, these results compare favorably to other speech intelligibility data from Vihman and Greenlee (1987) and Ware (1996). In Table 7, the available descriptive statistics for speech intelligibility are presented for the ages 3:0 to 4:0.The developmental progressions in the means and ranges of intelligibility from these studies are strikingly compatible. For the 3-year-olds (Vihman & Greenlee, 1987), the mean rate of speech intelligibility was 73%. Three raters were used to determine intelligibility, but not further methodological detail was reported. For children aged 3:6 (Ware, 1996), the mean was 88%. In Ware's study, two second year graduate students in speech and hearing sciences were used as listeners. All differences between the two sets of transcriptions were deemed significant. Differences among articles (*a*, *an*, and *the*) and plural markers, for example, were considered to be significant for intelligibility purposes in Ware's study. Ware reported that these two areas were a source of frequent disagreement in her transcripts. For the current study of children aged 4:0, the mean was 94%. This mean reflects several methodological differences in comparison to Ware, such as using 2 out of 3 listener agreement to determine intelligibility and excluding trivial differences from the error count such as article, number, and suffix differences. The current study considers certain small differences insignificant, therefore the intelligibility ratings are probably higher than would be true using Ware's methods.

As one would expect, the ranges of speech intelligibility among the three studies partially overlap: the 3:0 range was 54%-80%, the 3:6 range was 76%-96%, and the 4:0

range was 81%-100%. The standard deviation for Vihman & Greenlee (1987) was not reported, but the standard deviation for Ware (1996) and the current study are the same, 5.7.

Table 7

Comparison of Three Speech Intelligibility Studies: Ages 3:0, 3:6, and 4:0

Study	Vihman & Greenlee (1987)	Ware (1996)	Current
Sample Size	10	13	14
Age	3 years	3:6, +/- 2	4:0, +/- 2
Mean	73%	88%	94%
SD	Unreported	5.7	5.7
Range	54%-80%	76%-96%	81%-100%

Profiles by Performance and Individual History

Articulation Errors. In these speech samples from normal 4-year-olds, the number of phoneme errors on the Fluharty and APP-R: Preschool also did not predict the ranking of children in overall speech intelligibility. The lowest rate of intelligibility, for example, was measured in a female subject who made 4 phoneme errors on the Fluharty and no phonological deviations as measured by the APP-R: Preschool. Conversely, the male with the good rate of overall intelligibility (98%) had the largest number of phoneme errors (9) on the Fluharty and one phonological deviation as

measured by the APP-R: Preschool.

Stimulus Materials. Another methodological issue related to the books and manner of speech elicitation. This study used two books: *The Relatives Came* and *Good Dog, Carl*. Of these two books, *Good Dog, Carl* was preferable for age 4:0, +/- 2 months. Generally, the sequence of activities depicted in *Good Dog, Carl* was more obvious to the children. *The Relatives Came* began with a car trip to visit the relatives and included several scenes with families interacting. For many children, this type of family reunion may not have been familiar. Many children simply stated that they did not understand what was going on.

Verbal Prompts. In this study, special emphasis was placed on limiting the permissible verbal prompts that could be given to the child during the speech sample. As was described in the methods section, the researcher asked global rather than specific questions, such as *what is happening here*, and *what happened next* instead of asking specific questions such as *what do you see here* or *does he see the girl?* The examiner also used many nonspecific comments such as *wow*, *look there*, *oh*, *no*, and so forth. These restrictions were in place to discourage the child from repeating part of the same prompt in a reply, e.g., *What is the boy seeing*? *He is seeing the dog.* The intent was to allow the children to say what they wanted without providing many verbal models. This approach was only partially successful because several children seemed to miss the verbal scaffolding that is customary in the traditional picture book-reading dyad. For example, a parent might begin a story like this:

> *Well, let's look at this book. Look, Sarah. There's an airplane. Where do you think it is going? Do you like the color?*

When adults look at picture books with children, the adults may be more active than the child, making helpful remarks, drawing attention to key elements in the story, and providing comments that enhance comprehension. When the adult does not interact in the customary manner, some children may feel confused and may withdraw from the storytelling.

Subject 4, a female of 4:0, was cooperative in the telling of one story but was possibly unhappy or confused that she was not given more verbal feedback. This impression came through her comments and finally through her noncompliance with the plan to talk about both books. An excerpt from her speech sample illustrates her response to the examiner's indirect elicitation.

> What's that? What is he climbing? What's that? I don't know. He's swimming in the fish tank. I don't know. He is getting the bread out and the baby, he spilled the milk . . . Turn the page. Turn, turn, turn. I don't know it. I don't know. Now what are you doing? . . . .

One can speculate that subject 4 might have been more content if the examiner had provided more conventional responses, that is, the adult speaking more than the child, and had freely answered some of her questions. Following the paradigm for this study, however, the examiner did not provide direct answers to her questions but tried to respond satisfactorily within the constraints of the elicitation model.

Subject 11, a boy of 4:0, may have responded to the lack of verbal scaffolding by reversing the adult and child's roles. He was a questioner. Unlike subject 4, subject 11 appeared comfortable with the style of interaction. He posed his own questions and then

gave some answers.

> *What is the baby doing? What is the baby doing there? What did the baby get do? What did, what happened to the baby? I don't know. Why is the baby in there now? They / getting into the refrigerator. What, what is that? Margarine, grapes, tea, milk. I don't know. What is that? That is what happened. See, why is the baby's face like that?*

Greater flexibility in the scope of acceptable prompts and responses would be desirable to provide verbal responses that meet the individual needs of different children. Some children appeared to have no anxiety or concern about what the examiner said or did not say. Others seemed insecure and probably would have benefitted from more conventional book-reading dialogue.

Although the sample size for this study was small, a comparison of these results with the previous work of Vihman and Greenlee (1987) and Ware (1996) provides a good beginning for the collection of normative data in speech intelligibility. A comparison of these results with other research (Table 8) suggests that there is a steady progression in speech intelligibility development rather than a sudden burst of improvement at the end of the third year. Speech development may end with full intelligibility at age 4 as Weiss and Lillywhite (1981) had suggested or it may continue developing, as this study has found, into the child's fifth year of life. Extensive conclusions cannot be drawn from a study having a small sample size, but it is interesting that the range and standard deviation for 4 year olds, +/- 2 months was nearly identical to that of children aged 3:6, +/- 2 months (Ware, 1996).

CHAPTER V

SUMMARY

While SLPs routinely measure and report speech intelligibility when assessing young children and older clients with speech impairments, full normative data have not been available for comparison purposes. When assessing a young child to determine if a child's communication abilities are at or below that of peers, one must first know what the normative standards are. Knowing the normal distribution for speech intelligibility at several ages would allow for more precise uses of the intelligibility information than is currently possible. Only a few available studies exist to allow tentative normative comparisons of speech intelligibility data (Vihman & Greenlee, 1987; Ware, 1996).

The goals of this pilot study were to collect normative data on normal 4-year-olds, +/- 2 months, and to test procedural aspects of eliciting speech and determining speech intelligibility with listeners. Fifteen subjects, between the ages of 3:10 and 4:2, were recruited in the greater Portland metro area. All subjects were screened for normal hearing, normal expressive and receptive language, and normal phonological / articulation development. All subjects spoke English in the home and were reportedly free of motor, neurological, or developmental disorders. During the initial screening appointment, all of the selected 15 subjects displayed normal behavior. No significant speech-language deviations were noted during the initial screening and interview with parents.

The investigator recorded 15 speech samples on high quality digital audiotape in a soundproof booth at the Portland State University Speech-Language and Hearing Clinic in Portland. The 15 samples were shortened to 100+ words and compiled on

separate listening tapes. Three experienced graduate clinicians in speech-language pathology listened to the tapes with varied listening order. Listeners wrote full orthographic transcriptions of approximately 120 words, using slash marks for unintelligible words.

Once all samples were transcribed, the investigator compared the 3 written samples, determined a starting point, and took 100 continuous words from each subject to represent each subject. A word was counted as intelligible when at least two out of three listeners transcribed it the same. One outlier percentage was removed from the data because the speech intelligibility percentage for that child was 28 percentage points lower than any other. The resulting sample size was 14. A speech intelligibility percentage was derived and compared to establish normative data for 4-year-olds, +/- 2 months. The mean intelligibility percentage for the 4-year-olds in this study was 94%, the median was 96%, and the standard deviation was 5.7.

Research Implications

The use of 3 listeners for reliability purposes was an apparent success in this normative study when taking 2 out of 3 listener responses in agreement to be adequate proof of intelligibility. If unanimous agreement among the three transcriptions had been required, the resulting mean would have been lower than both the empirical and theoretical studies would have anticipated. The range and standard deviations would have been much larger. Using two out of three in agreement to establish intelligible responses appeared to equalize some individual listening and transcribing lapses in the group, such as overlooking quickly articulated portions or marking an excessive number

of unintelligible slashes in incomprehensible sections.

The approaches used to recruit subjects in future research should be more diverse. Many parents of young children cannot be reached by contacting preschools or posting notices in a newspaper. No one approach should be used to contact parents, but an assortment of approaches might result in a more evenly balanced demographic representation. Establishing contacts with pediatric health care professionals, various churches, synagogues, and community centers representing different socioeconomic groups would be very beneficial.

One hypothesis concerning respondents to this study was that only moderate-to-well educated parents were able to recognize the benefit of participation. Although free speech, language, and hearing screening was provided at the parents' preferred location, the response was very limited. Comments by participating parents showed good general awareness of the importance of early intervention. Many parents who did not avail themselves of these free services were probably uncertain as to the value of screenings in general. If a larger intelligibility study were undertaken, the researchers might first attempt to educate parents about the value of early screening and intervention before actually attempting to obtain subjects. Further research should be undertaken to learn what the general public believes speech, hearing, or language screenings accomplish. From those results, professionals in speech-language pathology and audiology might make better assumptions about the public that they serve and provide more meaningful information to the public to clarify misperceptions.

Future studies may want to establish clearer criteria for acceptable voice quality.

Common voice disorders that often are noted in children are vocal abuse or misuse of the vocal mechanism. Further research should be undertaken to determine the impact of the most common voice disorders on speech intelligibility of normal children. Research must also establish whether unusual voice characteristics (huskiness, high pitch) that are *not* of clinical concern to an SLP might have a disproportionately negative impact on intelligibility. Ultimately, a subject should be excluded from a normative sample if the impression of abnormality in a voice screening is sufficiently strong. In this study, no subject was excluded because of voice abnormality, but the criterion for acceptance was broad. If research determines that voice disorders or distinctive voice characteristics in the normal range have a significant impact on speech intelligibility, the screening methods employed in a normative study would need to be adjusted accordingly.

In this research, the transcripts reveal that several asides made by the children (e.g., *Could I touch the microphone? or Where is the stuffed bunny?*) were possibly more difficult to catch or decipher than more contextualized statements relating to the storybooks. Research has not established how much speech *without* a known context can be understood by an adult listener. The possibility exists that decontextualized comments may be unintelligible for many children of 4 years.

Future research must also resolve methodological questions such as how to best elicit speech, and use of stimulus materials are most appropriate for the child's age, cultural expectations, and interests or experience. While some issues in test administration cannot be resolved easily and will likely remain in the domain of the practitioner, SLPs may still need recommendations how one can elicit speech without

leading and influencing responses. Closer examination of the narrative structure of certain picture book stimuli would be useful. A reasonable goal might be to match the expected narrative maturity of a particular age group with materials that are at or somewhat below the anticipated levels. As with other assessment tools, the practitioners must focus on eliciting the best performance from a child lest the poor results be misinterpreted to the child's detriment. An SLP may be able in the future to use objective intelligibility measurement with more descriptive precision, but as with other forms of testing, will need to interpret findings with caution and possibly use multiple samples when in doubt.

Another area that may warrant further research is the relationship between narrative maturity and speech intelligibility. In the original sample of 15 children, one was very unintelligible and was noticeably poor in maintaining a coherent, logical narrative. While past research has looked into the relationships between various linguistic domains and speech intelligibility, narrative development may play an important role as it lays the foundation for comprehensible speech.

Clinical Implications

When SLPs try to distinguish between normal and delayed development in speech and language at a particular age, knowing the actual range of normal speech intelligibility development would allow objective ratings to be used more effectively in identification. Based on these preliminary normative results for children 4 years, +/- 2 months, it appears that SLPs should still anticipate some speech intelligibility variation at 4:0. Although this sample was small, the range of 81%- 100% provides a beginning

estimate of normal 4-year-old speech intelligibility that must be tested with more research. Considering that this sample was taken from children of middle class parents having a high school education or better, one may infer that research drawing from a more diverse sample would yield larger variation.

One implication from this study is that normal children may not be fully intelligible at 4:0 as Weiss and Lillywhite (1981) predicted. Generally, SLPs are concerned that a child has a communication impairment when a speech or language rating is at least 1.5 standard deviations below the mean. The mean derived from this research for children 4:0, +/- 2 months is 96%. With a standard deviation of 5.7, any speech intelligibility rating below 87% would indicate a possible disorder. Table 8 displays percentages 1.5 and 2 SD below the mean.

Table 8

Possible Standards for Identification of Communication Disorder in 4-Year-Olds Based on Speech Intelligibility Ratings: 1.5 SD and 2 Sds

Common Standards for Identifying Subnormal Communication Performance	Percentage Scores Signifying a Possible Communication Disability
1.5 Standard Deviations Below Sample Mean	87% and below
2 Standard Deviation Below Sample Mean	83% and below

Some medical and educational institutions and practitioners prefer to identify disabilities

by standardized scores that are two full standard deviations below the mean. Using present research as a preliminary standard, speech intelligibility scores below 83% should be of concern at age 4. All recommendations are presented cautiously due to the very small sample size and the likelihood that the results are demographically skewed. Higher intelligibility scores are probably seen in this sample than would be found in the true population. Considering these sample characteristics, using two SDs as the cutoff between normal and potentially abnormal speech intelligibility development is probably the most reasonable recommendation. Further normative data are needed for all ages of children between 3:0 and 5:0. SLPs still need to know the normal distributions, rates of change from age to age, and when most children can be expected to be fully comprehensible. Many normative studies are needed at key ages along the developing speech continuum: 3:0, 3:6, 4:0, 4:6, and 5:0. The current research data are so limited that the actual levels of intelligibility for normal children in this age range are still speculative. .

As previous research has shown, intelligible speech is the final product of a complex system. While articulation accuracy and overall phonological development is one predictor of overall speech intelligibility, several other factors including narrative ability and voice will have a strong effect. When better normative data for the development of speech intelligibility becomes available, educators, SLPs, and pediatricians may have a powerful tool for better describing the abilities of both normal and abnormal populations. A worthy speech-language screening instrument should be easily administered, should have standardized scores or normative data available for

comparison, and should capture data related to authentic communication acts. Potentially, speech intelligibility measurement could meet these criteria very well.

References

Adams, M. R. (1990). The demands and capacities model I: Theoretical elaborations. *Journal of Fluency Disorders, 15,* 135-141.

American Speech-Language-Hearing Association (ASHA) (1993). Guidelines for audiological services in the schools. *Asha, 30,* 61-64.

Applebee, A. (1978). *The child's concept of a story: Ages 2 to 17.* Chicago: University of Chicago Press.

Baker, L., Cantwell, D. P. & Mattison, R. F. (1980). Behavior problems in children with pure speech disorders and in children with combined speech and language disorders. *Journal of Abnormal Child Psychology, 8,* 245-256.

Berko, J. (1958). The child's learning of English morphology. *Word, 14,* 150-177.

Bernthal, J. E., & Bankson, N. W. (1993). *Articulation and phonological disorders.* Englewood Cliffs, NJ: Prentice Hall.

Brown, R. (1973). *A first language, the early stages.* Cambridge, MA: Harvard University Press.

Campbell, T. F., & Shriberg, L. D. (1982). Associations among pragmatic functions, linguistic stress, and natural phonological processes in speech-delayed children. *Journal of Speech and Hearing Research, 25,* 547-533.

Carrow, E. (1985). *Test of Auditory Comprehension of Language.* Austin, TX: Learning Concepts.

Colton, R. H., & Casper, J. K. (1996). *Understanding voice problems: A*

physiological perspective for diagnosis and treatment. Baltimore, MD: Williams & Wilkins.

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

Cornett, B. S., & Chabon, S. S. (1988). *The clinical practice of speech-language pathology*. Columbus, OH: Merrill Publishing Company.

Day, A. (1985). *Good dog, Carl*. Hong Kong: Green Tiger Press.

Flipsen, P. (1995). Speaker-listener familiarity: Parents as judges of delayed speech intelligibility. *Journal of Communication Disorders, 28*, 3-19.

Fluharty, N. B. (1978). *Fluharty Preschool Speech and Language Screening Test*. Allen, TX: DLM Teaching Resources.

Gaines, N. D., Runyan, C. M. & Meyers, S. C. (1991). A comparison of young stutters' fluent versus stuttered utterances on measures of length and complexity. *Journal of Speech and Hearing Research, 34*, 37-42.

Goehl, H., & Martin, P. (1987, November). Intelligibility and familiarity: Is experience enough? Paper presented at the Annual Convention of the American Speech-Language-Hearing Association, New Orleans, LA.

Gordon-Brannan, M. E. (1994). Assessing intelligibility: Children's expressive phonologies. (1994). *Topics in Language Disorders, 14* (2), 17-25.

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

Grice, H. (1989). *Studies in the way of words.* Cambridge, MA: Harvard University Press.

Grunwell, P. (1992). Principled decision making in the remediation of children with phonological disability. In P. Fletcher & D. Hall (Eds.), *Specific speech and language disorders in children: Correlates, characteristics and outcomes* (pp. 215-240). San Diego, CA: Singular Publishing Group.

Hodson, B. W. (1986). *The Assessment of Phonological Processes - Revised: Phonological Screening - Preschool.* Danville, IL: The Interstate Printers and Publishers, Inc.

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.

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

Huggins, A. W. (1977). Speech timing and intelligibility. In J. Requin (Ed.) *Attention and performance* (pp. 279-298). Hillsdale, NJ: Lawrence Erlbaum.

Ingram, E. (1976). *Phonological disability in children: Studies in disorders of communication.* London: Whurr Publishers.

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

Kelly, C., Dancer, J., & Bradley, R. (1986). Correlation of SPINE test scores to judges' ratings of speech intelligibility in hearing impaired children. *Volta Review, 88,*

145-150.

Kent, R. D. (Ed.). (1992). *Intelligibility in speech disorders*. Philadelphia, PA: John Benjamins Publishing Company.

Kwiatkowski, J., & Shriberg, L. D. (1992). Intelligibility assessment in developmental phonological disorders: Accuracy of caregiver gloss. *Journal of Speech and Hearing Research, 35*, 1095-1104.

Martin, E. A. (Ed.). (1981). *The Bantam Medical Dictionary*. New York, NY: Bantam Books, Inc.

Monsen, R. (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 speech intelligibility measure. *American Journal of Speech-Language Pathology, 4*, 22-28.

Morrison, J., & Shriberg, L.D. (1992). Articulation testing versus conversational speech sampling. *Journal of Speech and Hearing Research, 35*, 259-273.

Nicolosi, L., Harryman, E., & Kresheck, J. (1983). *Terminology of communication disorders: Speech-language-hearing*. Baltimore: Williams & Wilkins.

Owens, R. E. (1992). *Language development: An introduction*. New York: Macmillan Publishing Company.

Panagos, J. M., & Prelock, P. (1982). Phonological constraints on the sentence productions of language-disordered children. *Journal of Speech and Hearing Research, 25*, 171-177.

Paul, R. (1995). *Language disorders from infancy through adolescence:*

Assessment and intervention. St.Louis, MO: Mosby-Year Book, Inc.

Paul, R., & Shriberg, L. (1982). Associations between phonology and syntax in speech-delayed children. *Journal of Speech and Hearing Research, 25*, 536-547.

Poole, I. (1934). Genetic development of articulation of consonant sounds in speech. *Elementary English Review, 11*, 159-161.

Prather, E., Hedrick, D., & Kern, C. (1975). Articulation development in children aged two to four years. *Journal of Speech and Hearing Disorders, 40*, 179-191.

Rosenbek, J. C., & LaPointe, L. L. (1985). The dysarthrias: Description, diagnosis, and treatment. In D. F. Johns (Ed.), *Clinical management of neurogenic communicative disorders* (p.97). Boston, MA: Allyn and Bacon.

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 (Ed.) *Intelligibility in speech disorders* (pp. 11-34). Amsterdam: John Benjamins Publishing Co.

Schiavetti, N., Metz, D. E., & Sitler, R. (1981). Construct validity of direct magnitude estimation and interval scaling of speech intelligibility: Evidence from a study of the hearing impaired. *Journal of Speech and Hearing Research, 24*, 441-445.

Shriberg, L. (1980). Developmental phonological disorders. In T. Hixon, L. Shriberg, & J. Saxman (Eds.), *Introduction to communication disorders.* (pp. 263-306).

Englewood Cliffs, NJ: Prentice-Hall.

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

Shriberg, L. D., Kwiatkowski, J., Best, S., Hengst, J., & Terselic-Weber, B. (1986). Characteristics of children with speech delays of unknown origin. *Journal of Speech and Hearing Disorders, 51*, 140-160.

Smit, A. B., Hand, L., Freilinger, J. F., Bernthal, J. E., & Bird, A. (1990). The Iowa-Nebraska norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders, 55*, 779-798.

Smith, C. R. (1975). Residual hearing and speech production in deaf children. *Journal of Speech and Hearing Research, 18*, 795-811.

Swank, L. (1994). Phonological coding abilities: Identification of impairments related to phonologically based reading problems. *Topics in Language Disorders, 14* (2), 56-71.

Templin, M. C. (1957). *Certain language skills in children*. Minneapolis: University of Minnesota Press.

Van Riper, C. (1954). *Speech correction: Principles and Methods*. Englewood Cliffs, N.J.: Prentice-Hall, Inc.

Vihman, M. M. (1988). Later phonological development. In J. Bernthal and N. Bankson (Eds.) *Articulation and Phonological Disorders* (pp. 110-144). Englewood Cliffs, NJ: Prentice Hall.

Vihman, M. M ., & Greenlee, M. (1987). Individual differences in phonological development: Ages one and three years. *Journal of Speech and Hearing Research, 30,* 503-521.

Ware, K. M. (1996). *A pilot study: Normative study on the intelligibility of 3 ½ -year-old children.* Unpublished master's thesis, Portland State University, Portland, Oregon.

Wallace, W. F. (1971). The wabbit yost his tock. *Mother's Manual,* 28-38.

Webster, P. E., & Plante, M. S. (1992). Effects of phonological impairment on word, syllable, and phoneme segmentation and reading. *Language, Speech, and Hearing Services in Schools, 23,* 176-182.

Weismer, G., & Martin, R. E. (1992). Acoustic and perceptual approaches to the study of intelligibility. In R. Kent (Ed.) *Intelligibility in speech disorders: Theory, measurement and management.*(pp.67-118). Philadelphia, PA: John Benjamins Publishing Company.

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

Weiss, C. E., & Lillywhite, H. (1981). *Communication Disorders.* 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. *Journal of Speech and Hearing Research, 35,* 1316-1332.

Wiig, E., & Semel, E. M. (1984). *Language assessment and intervention for the*

learning disabled. Columbus, OH: Bell & Howell.

Wolfram, W., & Fasold, R. W. (1974). Social dialects as a field of inquiry. In The study of social dialects in American English. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Yoder, P. J., & Davies, B.(1990). Do parental questions and topic continuations elicit replies from developmentally delayed children: A sequential analysis. Journal of Speech and Hearing Research, 33, 563-573.

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

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

Yorkston, K. M., & Beukelman, D. R. (1981). *Assessment of intelligibility of dysarthric speech*. Tigard, OR: CC Publications.

Yorkston, K. M., Beukelman, D. R., & Bell, K. R. (1988). *Clinical management of dysarthric speakers*. Boston: Little, Brown and Company.

Appendix A

Advertisement for the Portland Parent newspaper

FREE SPEECH, LANGUAGE, & HEARING SCREENING

for nearly 4 or just 4-year-olds for participating in a PSU speech development study.

Call Jane Firestone at 977-0914 for details.

Appendix B

Informed Consent

Yes, I will allow my child to participate in this research project concerning speech development at the preschool site and at the PSU Speech-Language and Hearing Clinic. I acknowledge that my child will receive a free speech, language, and hearing screening as a benefit of participation. I also acknowledge the possibility that my child may find sitting in the soundproof booth strange or frightening.

Parent / Guardian Name: [print, please] Signature

_____ _____

Child's name: _____

Child's Birthrate: _____ My child is _____ years and _____ months

Day phone: _____ Evening phone: _____

Jane Firestone (503) 977-0914, E-mail: fire@cyberhighway.net

This thesis proposal was approved by the Office of Graduate Studies and Research at Portland State University: (503) 725-8410

Appendix C

Questionnaire

Child's Name _____ Birth date _____
Parent(s) _____
Address _____
Parent #1 - Highest Level of Education Completed / Current Occupation
_____/ _____
Parent #2 - Highest Level of Education Completed / Current Occupation
_____/ _____
Relationship of person completing the questionnaire to the child:
_____

1. Has your child ever been diagnosed as demonstrating any of the following:
   Neurological impairment ? Yes ___ No ___
   Orthopedic or physical handicap? Yes ___ No ___
   Motor or movement impairment? Yes ___ No ___
2. Has your child had a history of ear infections as indicated by the following:
   Complaints of ear aches? Yes ___ No ___
   Actual infections? Yes ___ No ___ When was the first? _____
   If yes, how many times? _____
   When was the last time? _____
   Has had medical treatment for ear aches? _____ How many times? _____
   When was the last treatment? ____
   Ventilation tubes ("tubes") inserted in the ear drum(s)? Yes ___ No ___
   If yes, when? _____
   Are tubes currently in one or both ears? Yes ____ No _____
3. Information about speech development:
   Is English the primary language spoken in your home? Yes ____ No ____
   When did your child say his or her first word? _____
   When did your child begin to put two words together? _____
   Do family members understand your child's speech? Yes _____ No ____
   Do persons outside the family understand your child's speech? Yes ___ No ___

Appendix D

Parent Information Letter

Dear Parent or Guardian:

Please read this letter if your child just turned 4 or is soon to be 4. I am a graduate student at Portland State University in Speech-Language Pathology. I am conducting some research in children's speech development. To learn how much of a normal 4-year-old's speech is understandable to adult listeners, I will be testing children between the ages of 3 years, 10 months and 4 years, 2 months. Under the direction of Dr. Mary Gordon-Brannan at PSU, I will analyze the speech of the children who participate and will determine what percentage of their speech is understandable. This percentage reflects the children's " speech intelligibility." These data will help future clinicians know when referring a child for speech services is appropriate. The current research does not provide a clear picture of what is normal for 4-year-olds.

Children who participate in the study will interact with me first at their school and later at the Portland State University Speech and Hearing Clinic. If you give permission in writing, your child will receive a free and complete hearing, speech, and language screening. Each screening will take from 20-30 minutes. All of the screening will take place at the preschool site. During the screenings, your child will be asked to respond to sounds, identify common objects and pictures, repeat sentences, and point to pictures. Tape recordings will be made during part of the screening. You will be asked to complete a brief questionnaire concerning medical history, family, and speech and language milestones. The identity of the children and questionnaire information will be kept confidential. The results of all screenings will be given to you in writing.

The second phase of the study will take place at the Portland State University Speech and Hearing Clinic for better sound recording. I will contact you to make an appointment that fits into your schedule. Your child will only be asked to go to the PSU Clinic if (1.) he or she passes the hearing, speech, and language screening, and (2.) your child agrees that a visit to the clinic is acceptable. Your child's participation is wanted, but is totally voluntary. At the Clinic, your child will look at age-appropriate picture books and discuss pictures with me while the conversation is recorded on audio tape.

Please sign the form below and return it to your preschool when possible. I will be scheduling appointments after the initial screening phase at the preschool has been completed. If you have any questions, please feel free to call me at (503) 977-0914. You may withdraw from this study anytime without questions even if written permission was given. Thank you for considering my request. I hope to hear from you.

Sincerely,
Jane Firestone
(503)977-0914
E-mail: fire@cyberhighway.net

This thesis proposal was approved by the Office of Graduate Studies and Research at Portland State University: (503) 725-8410

Appendix E

Instructions for Listening

Thank you for participating as a listener in the thesis describing normal intelligibility at age 4. Please read the following directions and ask for clarification where needed. All listeners should follow the same protocol for consistency. You will receive 3 tapes labeled A, B and C. You will be assigned these tapes in a particular order. Once you finish transcribing the samples on the first tape, you may return the tape and transcriptions to Jane and obtain the next set. Plan to allot time on 3 separate days for this project: review only one tape a day.

I. BEFORE YOU START

All listeners must pass a hearing screening test at 20 dB HL. Jane will arrange to do the screening at your convenience.

II. CHECKING OUT A TAPE AND THE DIGITAL TAPE RECORDER

Over the course of this listening assignment, you will need to check out 3 separate tapes and listen to them on the PCM-2300 Sony DAT recorder digital recorder available in the Speech Lab. Each tape must be reviewed on a separate day. Listening and transcribing may require between 1 ½ - 2 hours.

III. REVIEWING EACH TAPE

Five separate recorded speech samples from different subjects will be on each tape. You will listen to a total of 3 tapes, for a total of 15 subjects (3x5). Each participant will have a reference number (1-15). At the top of each transcription, be sure to write the subject's number. The names of the subjects will not be shared and must be kept confidential. Paper and pencils for the transcriptions will be supplied in a manila envelope.

Each child was asked to look at and describe the same 2 picture books, *Good Dog, Carl* and *The Relatives Came*. You, the listener, will need to review the books before starting the listening. transcriptions. You will be permitted to look at the picture books while transcribing the samples. You will receive thus receive approximately the same context clues as were present in the original taping.

You will listen to each sample 3 times and to write exactly what you hear. You will not be required to write any special notes or symbols.

- Rewind each tape before starting and reset the digital counter to zero.
- Look at the numbers at the top right corner of each of the 15 blank transcription sheets. These numbers represent start and end points on the digital tape for each of the 5 speech samples.
- Begin your transcription of the child after Jane's comment to the child that is written in bold print at the top of your first blank transcription sheet. You may want to transcribe Jane's statements on the transcript as points of reference. Put them in parentheses so they can be set apart from the child's words, e.g. (Jane: What is happening here?)
- Put a slash mark (/) for each word you think you missed or could not understand.
- For words you think might have understood, simply write your best guess. Question marks will be disregarded.
- Please use regular spelling. Do not use IPA.
- Double space your transcriptions and leave plenty of room.

IV. FINISHING

Make an arrangement with Jane about where to exchange tapes and leave transcripts when one is finished. Call or leave messages for her at 977-0914 if there are any difficulties. Your efforts are very much appreciated. Monday, Wednesday, and Friday, I will be available for calls during the day at 916-6437.

Appendix F

Intelligibility Order with Results from Questionnaire and Screening

Order by Phonological Intelligibility Deviations: Milestones	Subject Number of Number & Phoneme Age	Percentage of Reported Speech-Intelligibility Language	Gender	Reported Ear Infections	Errors
1st	12 / 3:10	100	M	3-4	0
1	delayed				
2nd	1 / 4:0	99	F	6+	0
2	WNL				
3rd	5 / 4:1	98	M	4	-velars
9	delayed				
4th / 5th	4 / 4:0	97	F	15+	0
0	WNL				
	15 / 3:11	97	F	4	0
1	WNL				
6th / 9th	3 / 4:0	96	M	0	0
3	WNL				
	10 / 4:0	96	M	0	0
3	WNL				
	11 / 4:0	96	M	6	0
0	WNL				
	14 / 3:11	96	F	often	0
5	WNL				
10th liquids	13 / 3:10 2	95 WNL	F	0	-
11th	5 / 3:11	94	F	0	0
0	WNL				
12th	7 / 4:1	86	M	3	0
0	WNL				
13th	9 / 4:2	85	M	1	0
0	WNL				
14th liquids	8 / 3:10 5	81 WNL	F	3-4	-
15th	3 / 4:0	53	F	4	0
4	WNL				

Appendix G

Reported Speech-Language Milestones

Subject Number	First Word Spoken	First Two-Word Combination
1	8 months	18 months
2	10-12 months	12-15 months
3	Before 12 months	After 12 months
4	5 months (ma ma)	12 months
5	20 months	24+ months
6	24 months	25 months
7	12 months	18 months
8	Before 24 months	12-24 months
9	12 months	13 months
10	9 months	9 months
11	Korean until 8 ½ months	Can't remember
12	24 months	30 months
13	12 months	18 months
14	18-24 months	24 months
15	12 months	Unknown

Appendix H

Methods of Rating Intelligibility with Three Listeners:
Two out of Three or Three out of Three in Agreement

Subject Number	Intelligibility Percentage: 3 out of 3	Intelligibility Percentage: 3 out of 3
1	99%	93%
2	96%	96%
3	53%	33%
4	97%	81%
5	94%	81%
6	98%	81%
7	86%	71%
8	81%	64%
9	85%	65%
10	96%	86%
11	96%	90%
12	100%	92%
13	95%	82%
14	96%	86%
15	96%	86%

Appendix I

An Example of Intelligibility Scoring: Comparing 3 Transcriptions of Subject 7

Listeners	**Words Transcribed**								
	I	**II**	**III**	**IV**	**V**	**VI**	**VII**	**VIII**	**IX**
1:	all	skipped	away	/	/	/	/	/	/
2:	-	/	away	Mommy	they	came	to	get	-
3:	all	skipped	away	/	and	/	/	/	-
Scores:	1	1	0	0	0	0	0	0	-

Note: In overall scores, 1= intelligible, 0 = unintelligible.
In individual transcriptions, a slash mark indicate listener judged word unintelligible, hyphen = extra word.