This is to certify that the

thesis entitled

PERFORMANCE-BY-INTENSITY FUNCTIONS
OF NORMAL HEARING CHILDREN
ON TWO MULTIPLE-CHOICE TYPE PICTURE TESTS
OF SPEECH DISCRIMINATION

presented by

WILLIAM RICHARDSON CULBERTSON III

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Audiology & Speech Sciences

Major professor

Date 6-30-81

O-7639

PERFORMANCE-BY-INTENSITY FUNCTIONS OF NORMAL HEARING CHILDREN ON TWO MULTIPLE-CHOICE TYPE PICTURE TESTS OF SPEECH DISCRIMINATION

By

William Richardson Culbertson III

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Audiology and Speech Sciences

1981

ABSTRACT

PERFORMANCE-BY-INTENSITY FUNCTIONS
OF NORMAL HEARING CHILDREN
ON TWO MULTIPLE-CHOICE TYPE PICTURE TESTS
OF SPEECH DISCRIMINATION

By

William Richardson Culbertson III

The speech discrimination performances of 40 five- and ten-year-old children were examined as functions of four equal increases in stimulus presentation level in the absence and presence of background noise. Forty additional subjects were tested at one specified sensation level. All subjects passed a speech articulation, receptive vocabulary and pure tone audiometric screening before being presented with tape-recorded versions of the WIPI and GFW, two multiple-choice, closed message set speech discrimination tests requiring no verbal response. Performance-by-intensity function curves for all subjects accelerated then plateaued. Curves for the five-year-old group with background noise plateaued at between 85 and 90 percent correct discrimination. The other curves plateaued at or near 100 percent. There was no rollover. The GFW and WIPI elicited equivalent performances when administered under the same conditions. Children who had immediate prior experience with each test achieved higher scores at a specified sensation level than children who received a single presentation at the same

level.

Conclusions indicated that children require quiet listening conditions and adequate stimulus intensity for maximum speech discrimination performance. The WIPI and the GFW may be used alternatively as measures of speech discrimination.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Chapter 1

INTRODUCTION

An individual's ability to distinguish one speech sound from another is of interest to professionals from several different fields. Among these could be counted the elementary school teacher and school psychologist, but, more specifically, the speech-language pathologist and audiologist. The presentation of one of several word lists, designed accoding to a model described by Egan (1948), can be used to assess speech sound recognition when the individual can reliably respond by repeating the stimulus item. The sound levels involved are carefully controlled and constantly monitored. When evaluating children, however, variables related to age or communicative abilities may affect results to the extent that the scores obtained are not valid. The need to accomodate children is great since they comprise a large portion of audiologists' and speech-language pathologists' caseloads. For example, at the Albemarle Regional Center for Communication Disorders, an eastern speech and hearing clinic, approximately three quarters of the audiologists' evaluations were performed on children between the ages of five and ten years.

Audiologists may choose to employ a multiple-choice picture test of auditory discrimination such as the <u>Word</u>

Intelligibility by Picture Identification (WIPI) (Ross and Lerman, 1971) to accomodate the limitations of children's expressive communicative abilities. The test was conceived as a means of assessing the speech discrimination ability of hearing impaired children. The authors of the WIPI felt that conventional speech discrimination tests were unsuitable for use with hearing impaired children because of the unfamiliarity of the stimuli and the inability of the hearing impaired children to make suitable responses. The WIPI was designed with stimuli which were within the "recognition vocabulary" of a group of hearing impaired children. The response criterion was met with a picture pointing identification task. The score is derived from the percent correct of the responses made by the tested individual following the auditory presentations (Ross and Lerman, 1971).

On the other hand, the Goldman-Fristoe-Woodcock Test of Auditory Discrimination (GFW) (Goldman, Fristoe, and Woodcock, 1971) is an instrument that is widely used by speech-language pathologists who are interested in the relative strengths and weaknesses of their clients who have speech articulation disorders. The GFW was designed to provide measures of speech-sound discrimination ability under ideal listening conditions plus a measure of speech-sound discrimination in the presence of controlled background noise. The authors stated that their test is useful to clinicians in the area of speech and hearing as well as those of school psychology, remedial reading, and

learning disabilities. The examiner scores the GFW by counting the errors of the tested individual and translating the error score into a standard score of a percentile score. Translation of error scores into standard or percentile scores is accomplished by referring to a matrix of normative data supplied by the authors (Goldman, Fristoe, and Woodcock, 1971). The responses of the individual to this test may be used by the clinician to develop a suitable speech articulation training program. Although the authors clearly stated in the test manual that the norms supplied with the test apply to subjects listening through earphones, administration of word discrimination tests in a sound field arrangement is common (Sanderson-Leepa and Rintelmann, 1976). This is especially true in locations with limited equipment or in situations where the client rejects wearing earphones. Consequently, one might expect to find discrepancies between two sets of discrimination test results when one is obtained under controlled conditions and the other is obtained under relatively uncontrolled conditions. The validity of the standard score is, therefore, compromised in some situations.

Purpose

Both the WIPI and the GFW share a basic design: the tested individual points to one of several pictures in response to a spoken carrier phrase and stimulus word. The discrimination score depends upon the number of correct pointing responses made. One way in which the tests differ is that the WIPI was designed to be administered through a speech audiometer at a specified hearing level, whereas the GFW authors state only that the test should be administered, preferably through earphones, with the volume set at a comfortable level of loudness. It should, therefore, be of interest to both speech pathologists and audiologists to study the extent to which these two widely used diagnostic measures compare and the extent to which noise affects the scores.

Specifically, there were four purposes: first, to compare the performances of two different age groups of normal hearing children as a function of intensity in the sound field on two multiple-choice type picture tests of auditory discrimination, the WIPI and the GFW; second, to identify an optimum intensity in decibels above the speech reception threshold at which these tests should be presented to children in the sound field represented by the sample herein; third, to identify the possible existence of a significant difference in performance on these tests between two age groups, five years and ten years, in the presence of background cafeteria noise; and fourth, to examine the effects of multiple presentations upon listener performance.

Importance

This study is important for the following reasons. First, little information is available on the performance of children as a function of intensity on multiple-choice type picture identification tests. Such information would be useful to the audiologist and the speech-language pathologist who cannot employ the conventional type of discrimination test with a given client because of the client's age or the presence of an expressive communication disorder that would spuriously affect the results. Second, the GFW is often presented to children at intensity levels which are unknown or unspecified by speech-language pathologists and other examiners. In many cases, educational program planning is based on these results. It could be argued that conclusions based on results obtained at arbitrary and unknown intensities could provide erroneous results because of a confounding variable. A third reason lies in the fact that the GFW and the WIPI scores may be commonly compared since the GFW is often used by speech-language pathologists with the intention of referring the tested individual to the audiologist for a more complete evaluation of receptive communicative abilities, and the audiologist frequently chooses the WIPI as a means of quantifying speech discrimination skill. It is important, therefore, to examine the similarity of children's performances on these two tests as a function of intensity. Finally, this study is needed to provide further comparison data among speech

discrimination word lists. A comparison of the GFW and WIPI might provide the audiologist with information which could lead to an alternative multiple-choice type picture discrimination test.

General Applicability

A comparison of the WIPI to the GFW may have import beyond the limits of the study itself. Since the children involved in this study have normal hearing as assessed by audiometric screening and normal receptive vocabularies as verified by scores obtained on the Peabody Picture Vocabulary Test (Dunn, 1965), it may be assumed that the data obtained are useful to the clinician using either the WIPI or the GFW with other children who are within the age ranges sampled. Further, the concept of a discrimination test in which the subject responds by pointing to one of a set of pictures is a useful one in audiological testing of populations who cannot respond to conventional speech identification testing. It is suggested that these data contribute to further research on multiple choice-type picture discrimination tests.

Finally, the means of generating the speech stimuli and the articulation curves described in this study should interest speech and hearing clinicians who might wish to develop a standard discrimination list presentation technique, based on recordings of their own voices.

Definition of Terms

For the purpose of this study, the following terms are defined to provide a common basis for understanding:

Speech or word discrimination: The ability to distinguish speech sounds from one another within the context of certain words formed with these sounds.

Auditory discrimination: The ability to distinguish one speech sound from another (Wood, 1957).

Multiple-choice-type picture discrimination test: A testing instrument which is used to assess auditory discrimination of speech sounds by quantifying the number of accurate picture pointing responses made by the subject. The subject is asked to point to one of several pictures which best represents the spoken auditory stimulus word. No verbal response is required of the subject. In the case of the GFW, there are four alternatives. In the case of the WIPI, there are six.

Performance-by-intensity function (P.I. function): A change in an individual's performance on a perceptual task brought about by an increase in the intensity of the stimulus.

Performance-by-intensity curve (P.I. curve): A graphic representation of the P.I. function with the subject's performance plotted on the axis of the ordinates and the stimulus intensity plotted on the axis of the abscissas.

Statement of Research Questions and Hypotheses

The primary purpose of this study was to compare the performances of children in two age groups on two multiple-choice type picture tests of speech discrimination, the GFW and the WIPI. The concept of a speech discrimination test which requires no oral language output from the individual being tested appears to be a useful one under certain circumstances. It is important, however, to investigate the effects of signal intensity on the performances of children taking these tests so that a more valid clinical interpretation of the results of these tests may be developed.

Research Questions

The following five questions represent the primary thrust of this research project:

1. Is there a difference between the mean performance-by-intensity functions obtained with the GFW and the WIPI? More specifically, will systematically increasing the intensity of the presentations of the stimuli result in differences between the slopes of the linear portions of the graphed mean P.I. function?

2. Does cafeteria background noise affect the performance on these two tests?

3. Is there a difference in performance on these two tests associated with the ages (five years and ten years) of the two groups of children tested?

4. To what extent can the scores obtained with the GFW presented at a specified sensation level be used to predict the scores obtained with the WIPI at the same sensation level?

5. Is there a difference in performance on these two tests associated with single versus multiple presentations of the stimulus materials?

Research Hypotheses

From the research questions for this study, the following hypotheses were drawn:

1. There is no more than 2%/dB difference between the slopes of the portions of the mean P.I. function curves between 8 and 16 dB SL for the two tests within each age group, in both quiet and noise listening conditions.

2. There is no difference in the minimum sensation levels required for all children within each age group to receive a score of 90% or better correct discrimination on both tests in the quiet listening condition.

3. There is greater than 2%/dB difference between the slopes of the portions of the mean P.I. function curves between 8 and 16 dB SL for both tests between age groups.

4. There is no difference in the mean minimum sensation level of 90% or better correct discrimination

for each subject in the quiet listening condition between age groups for both tests.

5. Correlations between scores on the two tests within each age group and at each sensation level are significant at the $p \leq .01$ level for both listening conditions.

6. Differences between discrimination scores related to age are significant at the $p \leq .01$ level at each sensation level.

7. Differences between mean discrimination scores related to listening condition are significant at the $p \leq .01$ level.

8. There is no difference ($p \leq .01$) between the mean speech discrimination scores of the experimental group that hears repeated presentations of the two tests and those of the group which hears only one presentation of the test at a sensation level at which the repeated presentation group first obtained a score of 90% or better.

Chapter 2

REVIEW OF RELATED LITERATURE

The review of literature presented in this chapter is divided into three main sections. The first section deals with literature related to auditory discrimination of speech sounds, the second section contains a review of research concerned with factors involved in the measurement of speech discrimination ability in human subjects. The final section involves an inspection of literature regarding the characteristics and use of the performance-by-intensity function.

Auditory Discrimination

The speech-language pathologist views speech discrimination from a slightly different perspective from that of the audiologist. The speech-language pathologist is concerned with the discrimination of speech sounds as one of several basic functions upon which the production of intelligible and acceptable speech is based. The audiologist, on the other hand, investigates speech discrimination as a means of determining the functional adequacy of the auditory mechanism and pathways as well as of electro-acoustic appliances prescribed to facilitate the receptive aspect of communication.

For many years, oscilloscopic and spectrographic studies of phoneme utterances have been accepted as showing that discernible differences in formant frequency and intensity exist between and among both vowels and consonants (Stevens and House, 1961; Tarnoczy, 1948). These differences are present whether the phonemes are uttered in isolation or in connected speech (House and Fairbanks, 1952; Miller and Nicely, 1954). Vowel sounds have been synthesized successfully with resonance types of speech synthesizers which require only a knowledge of spectral energy concentrations or formant frequencies (Flanagan, 1957). Differential perception of phoneme categories is assumed to be based upon variations in these physical parameters (Mattingly, Liberman, Syrdal, and Hawles, 1971).

As defined by the American Heritage Dictionary of the English Language (Morris, 1971), the word "auditory" refers to that which pertains " . . . to the sense, the organs, or the experience of hearing," whereas the word "discrimination" means "to perceive the distinguishing features of; recognize as distinct." Combining these two general definitions, it would appear that "auditory discrimination" refers to the act of perceiving the distinguishing features among some forms of acoustic energy.

Wood (1957) defined auditory discrimination as the "ability to discriminate between sounds of different frequency, intensity and pressure pattern components; the ability to distinguish one sound from another." This

definition is sufficiently broad to account for a literal interpretation of the term, as long as specific alternative responses among which to discriminate are available. Wood's definition is perhaps too broad to be used to specify that component of human behavior that is being studied herein.

The speech-language pathologist's view of auditory discrimination has been represented by Wepman (1958), Weiner (1967) and Muma (1978). Wepman implied that auditory discrimination is the " . . . ability to recognize the fine differences that exist between the phonemes that are used in English speech" in the manual of administration, scoring and interpretation for his Auditory Discrimination Test (1958, 1973). It would appear that misunderstanding could result if auditory discrimination is seen as an individual's ability to recognize "fine differences," unless these differences were specified in terms of the acoustic parameters mentioned above. Instead, it might have been more appropriate for Wepman to have used a term which described the ability to recognize some phonemes used in English words.

Weiner (1967) has referred to auditory discrimination as the "ability to distinguish between closely related speech sounds." This description is similar to Wood's (1957) "ability to discriminate between sounds of different frequency, intensity, and pressure pattern components."

Muma (1978) regarded the differential perception of speech sounds as "the categorization of phonemic patterns

and intonations." He held that the type of behavior exhibited by young children in which differential responses to speech stimuli are observed is not merely discrimination, but categorization of phonemic classes. Muma saw perceptual context as essential in the discrimination/categorization component of early language development.

Audiologists generally view the perception of phoneme differences as one aspect of overall auditory ability. Davis (1970), Sanders (1971), Newby (1974), Tillman and Olsen (1974), and Goetzinger (1972) have discussed phoneme perception in terms that are representative of the audiologist's point of view.

Davis (1970) described speech discrimination as one facet of auditory perception. His discussion of speech discrimination testing refers to speech discrimination testing as "articulation testing." As Davis explained, the term "articulation," applied in this manner, was originally used by telephone engineers to describe the adequacy of a telephone system in transmitting speech signals. Sanders (1971), when discussing aural habilitation, differentiated between gross sound discrimination and speech sound discrimination as components of the general ability of auditory discrimination. Sanders defined auditory discrimination as " . . . using the characteristics of a sound to differentiate between various auditory experiences." Newby (1974) distinguished between "gross sound" discrimination, such as the ability to distinguish between the sounds of a bell and

a horn; "gross speech" discrimination, such as the ability to differentiate between dissimilar speech sounds (e.g., /a/ and /ɛ/); and "fine speech" discrimination, such as the ability to discriminate between spectrally similar sounds (e.g., /f/ and /θ/).

Tillman and Olsen (1974), in their discussion of speech audiometry, differed somewhat from their audiologist colleagues by using the terms "auditory discrimination," "speech intelligibility," and "speech discrimination" interchangeably. Speech "recognition" is an appropriate term when the range of alternative responses is infinite, as is the case with an open-set test paradigm.

The term "word discrimination," as used by Goetzinger (1972), is most suitable to describe that hearing function which is the focus of the present investigation. The speech sounds are couched in words, rather than being presented in nonsense syllables or in isolation. Further, specific alternatives are presented in a closed-set paradigm. This is an important distinction, since the primary purpose of word discrimination testing is to determine how well a person is able to perceive accurately the meaning of spoken words in a society in which the basic mode of communication is through speech (Goetzinger, 1972). The evaluative instruments examined here purport to require an individual to recognize one word as being distinct from several others in a closed-set paradigm.

Development of Speech Discrimination Abilities

To understand the need for gathering data on the results of word discrimination testing with different age groups, a review of the literature dealing with developmental trends in auditory discrimination is helpful. Here, research will be reviewed which has studied discrimination ability as it develops from infancy through the school-aged years and finally into adulthood.

Infant speech discrimination. It has been demonstrated that infants appear to enter the world with at least some sensitivity to the phonological structure around them. Specifically, infants have given differential responses to natural and synthetic speech stimuli which contain acoustic cues which are sufficient to elicit differential perception of phoneme categories in adults. The recurring theme in the literature on infant speech perception is the distinction between auditory versus phonetic discriminative abilities. The term "categorical perception" is applied to the phenomenon which occurs when listeners' discrimination of two stimuli is limited to those stimuli that they can label or identify differentially (Morse, 1978). The basis for studying infants' discrimination of speech has been research involving adult listeners wherein they demonstrated better between-category than within-category discrimination. In one such study, the adult subjects demonstrated that changes in the initial transition of the second formant (F2) of

synthetic speech stimuli are sufficient to cue the distinction between the stop consonants /b/, /d/ and /g/ (Mattingly, Liberman, Syrdal and Hawles, 1971). This information was used to develop synthetic speech stimuli categories for infants studies (Morse, 1978).

Research related to the categorical perception of speech by infants has relied primarily on two paradigms: an operant nonnutritive sucking procedure and a heart-rate orienting response design. In the first paradigm, a speech signal was delivered to the infant contingent on the infant's rate of vigorous sucking. This hard sucking was referred to as high amplitude sucking (HAS). The amplitude of the sucking was monitored through pressure transducers contained in an artificial nipple. The infant's rate of sucking (per minute) increased from the baseline rate of approximately 20-30 responses per minute at a specified amplitude as he or she discovered the contingent relationship of the speech signal to the high amplitude sucking. This increase was termed "acquisition." It was presumed that the infant increased his or her rate of sucking in response to behavior reinforcing properties of the stimulus. The infant began to satiate, or habituate, the HAS response after continued presentation of the speech signal. Then the rate of sucking decreased. If the rate of decrease met a predetermined criterion, a new stimulus was presented to the experimental group, while a control group continuously received the original stimulus. Discrimination of the stimuli was

inferred if there was a significant increase in mean HAS rate among the experimental group relative to the rate among the control group and coincidental to the stimulus change (Eimas, 1974a). The HAS paradigm was used to demonstrate that infants can perceive differences between consonant-vowel syllables. Trehub and Rabinovitch (1972), for example, used the HAS design to conclude that babies as young as four weeks were able to perceive differences between synthetic versions of /b/ and /p/ as well as between natural speech versions of /b/ and /p/ bilabial stops and /t/ and /d/ tip-alveolar stops.

The heart-rate orienting paradigm (HR) utilizes cardiac deceleration, an orienting response, to mark an infant's awareness of a stimulus change. Berg (1971) used the HR design to conclude that infants perceived temporal and frequency changes in pure tone stimuli. Berg observed rapid cardiac deceleration following changes in 1100 Hz. and 1900 Hz. tones.

Infant studies were oriented toward an answer to the question of whether the infant's discrimination of speech reflects any sensitivity to phonetic categories used by adults or whether the differential responses observed were merely indicative of simple auditory discriminative abilities. In one of the earliest studies of infant speech perception in five to six month olds Moffitt (1971), using the HR paradigm, observed apparent categorical perception of second formant (F2) transitions in three-formant, synthesized /ba/ and /ga/ syllables. The F2 transitions started at 846 and 2078 Hz., respectively. Both F2 transitions approached a steady state value of 1075 Hz. during a time interval of .055 seconds.

Later, Morse (1972) obtained similar results with an HR paradigm, showing categorical perception of second formant transitions in synthetic speech. He later (1978) stated that the results do not imply discrimination of phonetic consonant categories but, more specifically, are indicative of frequency discrimination by the subjects.

Eimas (1974a) used an HAS paradigm to study phonetic perception among two- to three-month-old infants, presenting subjects with either a between-category discrimination (e.g., /dæ/ versus /gæ/), a within-category discrimination (e.g., within /dæ/), and a control group having no shift. Infants discriminated only on the between-category condition, leading Eimas to infer adult-like phonetic discrimination among the subjects. The findings of Eimas (1974a), Miller and Morse (1976) and Till (1976) suggest that infants discriminate differences in place of articulation not only auditorily but phonetically as well.

Additional studies of infant speech discrimination have supported conclusions that infants exhibit phonetic discrimination of voice onset time (Eimas, Siqueland, Juszyck and Vigorito, 1971; Trehub and Rabinovitch, 1972) as well as categorical discrimination of "liquid" consonants (/ra/ versus /la/) (Eimas, 1975a). Eilers and Minifie

(1975) obtained results which support infants' discrimination of fricative cues for /sa/ versus /va/, /sa/ versus /ʃa/, but not /sa/ versus /za/. This apparent contradiction to the Eimas et al. (1971) study was explained in terms of the possible masking effects of the high frequency friction noise source over the lower frequency F1 and F0 cues (Morse, 1978). Stevens and Klatt (1974) posited that the results of the Eimas et al. (1971) investigation indicated that one- and four-month-old infants cannot generally discriminate such small (.020 seconds) differences in voice onset time and that the discrimination evidenced by the between-category shift group (e.g., .020 seconds versus .040 seconds) may evidence the infant's ability to discriminate the presence versus the absence of an F1 transition. Miller, Morse and Dorman (1977) have observed discrimination of initial burst cues (.008 to .032 seconds in duration) for the consonants /bu/ and /gu/.

The validity of the HAS paradigm has been questioned by Butterfield and Cairns (1974). Their criticisms are based on what they view as a lack of control groups in the research design. These control groups, Butterfield and Cairns argue, should include one group which receives no auditory stimulation and another which receives non-contingent auditory stimulation. Following this reasoning, it appears that a relationship between the presentation of a stimulus and the pattern of sucking responses cannot be demonstrated until differences between experimental and adequate control groups can be demonstrated as non-significant. Butterfield and Cairns (1974) compared the responses of three groups of infants: a shift (in stimulus) group, a shift-to-silence group. and a no-shift group. They felt that their data provided limited support for the postshift rationale of the HAS procedure. King (1976) demonstrated the possibility that the pattern of results shown by a shift group could be occuring as a random event rather than as a result of VOT discrimination and stated that Butterfield and Cairns failed to employ an adequate control group. King's remarks may stem from the fact that Butterfield and Cairns combined the frequencies of responses of the shift-to-silence and the no-shift groups since the data for the two groups were so similar.

Other criticisms of the HAS procedure include questions regarding the reinforcing properties of the auditory stimuli as well as concerns over the paucity of within-category stimuli (Butterfield and Cairns, 1974). Further, the fact that synthetic speech stimuli are complex, as pointed out by Stevens and Klatt (1974), tends to modify conclusions based on these data. Other acoustic features comprising the synthetic speech stimuli may actually be what is being discriminated.

Eilers, Wilson, and Moore (1977) used a visually reinforced infant speech discrimination paradigm. This operant design made use of a head-turning response reinforced by a lighted toy animal. This approach was seen as advantageous because the reinforcer was independent of the stimulus. In the HAS studies, the stimuli are both reinforcers and discriminative stimuli. Eilers et al. felt that for this reason it was appropriate to use their visual reinforcement paradigm after the subjects reached four months of age. Their data, obtained using naturally produced, tape recorded speech as stimuli, suggested that infants six- to fourteen months old can be tested for perception of subtle speech contrasts. The stimuli consisted of voiced and unvoiced plosive and fricative consonant sounds.

An interesting line of thought was pursued regarding infants' abilities to discriminate voice onset time (VOT). Cross-language studies were primarily oriented to examining the perception of VOT by infants using languages which have three voicing categories. Whereas English has two voicing categories (e.g., voiced and unvoiced) which contrast phonemically, other languages may distinguish between a pre-voiced category (e.g., Lebanese, Arabic, Spanish) and a voiced category. Thai employs all three categories, including voiceless, prevoiced and voiced (Lisker and Abramson, 1964). Cross-language VOT studies were undertaken to determine whether discrimination of voicing categories requires any previous language experience or whether it is an innate ability (Lasky, Syrdal-Lasky and Klein, 1975; Streeter, 1976). These cross-language studies do

not tend to support the hypothesis that infants can auditorily detect certain category boundaries not available in their native languages. Statistical significance regarding the categorical nature of this discrimination has not yet been achieved.

The perception of vowels by infants has been studied using the HAS paradigm. Trehub (1973b) found that infants can discriminate between isolated vowels (e.g., /i/, /a/ and /u/) as well as vowels in a consonant vowel syllable context (e.g., /pa/pi/, /ta/ti/ and /pa/pu/). Swoboda, Morse and Leavitt (1976) found that eight-week-old infants tend to perceive vowels continuously: discrimination occurred equally as well within- and between-categories.

To summarize the findings of research related to infant perception, it appears that infants as young as the neonatal stage may be able to discriminate between speech sounds in a manner similar to adults. Findings are encouraging when the stimuli have been differentiated by place as well as manner of articulation or by voice onset time. In addition, subjects appeared to discriminate vowels, both in isolation and in consonant-vowel (CV) syllables. The experimental designs used in these studies are under attack on several fronts, primarily regarding the lack of satisfactory controls and the necessarily complex nature of the stimuli involved.

Speech discrimination in children. The literature related to developmental trends in auditory discrimination abilities among older, school-aged children is not nearly so plentiful as is that related to infant speech perception. Many of the studies which use subjects in this age group are primarily oriented toward developing tests of word or speech sound discrimination abilities. In the studies to be reviewed here, it must be borne in mind that the ability to respond to the testing, to comprehend the instructions and to make a satisfactory response, is likely to have involved several other developmental factors in addition to the discrimination response (Mills, 1975; Schwartz and Goldstein, 1974).

Weiner (1967) reviewed sixteen studies concerned with the relationship of auditory discrimination skills and speech articulation. A major point of his discussion was that the ability to distinguish between speech sounds appears to be an age-related variable. Weiner noted the relationship between age and speech sound discrimination skills seemed to exist in children up to age eight.

Earlier, Templin (1943) found that errors on her own test of speech sound discrimination dropped regularly with age and performance approached a ceiling at about nine years. She noted that the position of the discriminative element within a syllable was important, with more errors being made when this element was in the medial position (intervocalic) or final position. She found that the position of the discriminative element was significant at the p=0.24 level among the second graders. Templin concluded that speech sound discrimination was a developmental phenomenon. However, there was a slight increase in errors between the fifth and sixth grade children, an increase attributed to the lack of emphasis on phonics in the sixth grade.

Templin's later study (1957) made use of the revised edition of her nonsense syllable test. She obtained similar results to her earlier study, noting that mean discrimination scores improved over the age range from six to nine years. The difference in mean scores was significant between ages six and seven (p=0.05); however, there was not a significant difference in mean scores between the seventh and eighth years. Children from more favored socioeconomic groups performed better than those from the less favored groups (p=0.01), and girls outperformed boys at the 0.05 level of significance. Templin noted in the latter study that her test, which demanded an understanding of the concepts of "same" and "different," required a considerable level of intellectual development. This problem was, of course, not dealt with by simply reducing the number of items used from seventy in the 1943 study to fifty in the 1957 study.

Wepman (1968) noted that discrimination ability increases or improves with age, maturing at least to the ninth year. His discrimination testing was similar in format to Templin's except that words were used instead of nonsense syllables. Later, Wepman (1973) reported an increase in median discrimination test scores at each age level fron ages five through eight (N=1000) in the revision of his Auditory Discrimination Test. Using Wepman's first test, Morency (1968) also found a significant difference ($p=0.01$) between auditory discrimination scores of children in the first and those of children in the third grades.

Mecham (1971) found that listening accuracy for words was a developmental phenomenon which rapidly attained a plateau between the fourth and sixth grades. He designed a multiple-choice type picture discrimination test for study by coupling words from the Thorndike and Lorge (1944) word list with picture materials from speech therapy. Children were presented three pictures and three stimulus words, only one of which matched one of the pictures. Each child had to couple the word with the picture by pointing to the correct picture. Although the results of this study are generally in agreement with data on the development of speech discrimination skills, the test paradigm used in Mecham's study seemed unnecessarily complicated, containing auditory as well as visual distractors. Certainly, there was an unnecessary amount of auditory memory involved in this type of format. Another shortcoming of Mecham's investigation was the apparently arbitrary control of stimulus intensity. The speech signals were delivered in free-field at "two-thirds" of the volume on a tape recorder.

Thompson (1963) tested the speech discrimination of 105 elementary school children before they entered the first grade, using the Wepman Auditory Discrimination Test (Wepman, 1958) and the Boston University Speech Sound Discrimination Picture Test (Pronovost and Dumbleton, 1953). The children were tested again after they completed the second grade. Findings indicated that the final testing scores were higher than the initial testing scores, although 100% performance was not attained by all the post-second graders. Here again, very little attention was paid to stimulus intensity control.

Sanderson-Leepa and Rintelmann (1976) compared the performances of children on three auditory discrimination tests; the WIPI, PBK-50 and the N.U. 6 tests. They used sixty normal hearing children in five age groups, from three and one-half to eleven and one-half years. Their data indicated a developmental increase in test performance accuracy as manifested by an increase in mean scores and a decrease in the standard deviations. Performance of the eleven and one-half year old group on the N.U. 6 was in close agreement with the findings of Tillman and Carhart (1966) as well as those of Rintelmann, Schumaier and Burchfield (1974) with normal hearing adults when stimuli were delivered at 32 dB SL.

Normative data of clinical tests of speech perception are consistent with the hypothesis that word discrimination is a developmental skill. Data supplied with the Goldman-Fristoe-Woodcock Test of Auditory Discrimination indicated that performance improved from about age four years, where the normative data began, to about twenty-five years. Performance declined after this point. These data also indicated that for children younger than about age ten, performance is affected by competing noise to a greater extent than that of older children and adults as old as eighty-six years.

In summary, research findings related to the development of speech discrimination skills in children during the school years have indicated a rapid increase in test performance as age increases up to approximately the ninth year. Some further improvement can be seen past this age. Conclusions based on these studies are tempered by the fact that it is difficult to separate concomitant developmental factors involved in testing.

Speech discrimination in adults. The relative paucity of studies concerned with the developmental aspects of speech discrimination among subjects who fall between the categories of elementary school-age and old age leads one to infer that little identifiable change occurs during those years.

As reported earlier, the standardization studies used to develop the Goldman-Fristoe-Woodcock Test of Auditory Discrimination have shown that the ability to distinguish between speech sounds continued to show some improvement past the popularly cited eight- or nine-year ceiling. This trend continued until age twenty to thirty, at which point mean performance levels began to decline. In these data, the oldest age group tested, sixty-eight to eighty-six years, performed at about the same level ($\bar{X}$ = 1.77; SD = 1.88) (Goldman et al., 1971).

The changes in discrimination ability that occur with old age are among the most noticeable changes which occur after the peak of word discrimination ability is reached. Oyer, Kapur and Deal (1976) described the hearing impairment that accompanies old age as " . . . devastating to the communication process." They estimated that the incidence of hearing loss in the older age group was approximately ten times as great as it is among those in early adulthood.

Gaeth (1948) described what he referred to as "phonemic regression." Certain older people had poorer performance in speech discrimination testing than would be predicted on the basis of other audiometric testing results. This phenomenon, seen most frequently in

individuals greater than fifty years old, has been observed by others (e.g., Dirks, 1978). It can be expected that as individuals develop the auditory problems associated with advancing age, some clinical manifestations will take the form of decreased performance on speech discrimination tests. It is noteworthy that many of Gaeth's subjects did not have difficulty with discrimination testing in spite of the presence of hearing losses.

Sensitized speech discrimination testing has shown changes in ability to be coincident with the increase in age among adults. For example, Konkle, Beasley and Bess (1977) presented the Northwestern University Auditory Test No. 6 to 118 subjects in four age groups ranging from 54 to 84 years (mean age = 68 years). Stimuli were presented at three sensation levels, 24, 32, and 40 dB., and at four stages of time compression, 0, 20, 40, and 60% of normal duration. Results indicated that intelligibility decreased as a function of increasing age and increasing percent of time compression. Conversely, intelligibility increased as a function of increasing sensation level.

The research reviewed in the above section relative to the development of speech discrimination skills has shown that speech discrimination abilities are testable after the onset of speech and show a steeply accelerated developmental improvement until the ninth year. After this point, improvement is less dramatic. After about age thirty, a decline in

performance is observed, exacerbated by the occurrence of auditory problems associated with aging.

Speech Discrimination as a Correlate of Other Speech and Language Skills

In this section, literature will be reviewed in which performances of individuals in the areas of speech articulation and language development are compared to their performances in response to speech discrimination testing. First, studies will be discussed which examine the relationship of speech articulation with discrimination. Following this, a review of data related to various language skills, as correlates of speech discrimination skills, will be presented. Language skills examined will include intelligence, psycholinguistic skills and dialect. Finally, factors related to the individual's environment will be examined.

Speech discrimination and articulation. Travis and Rasmus (1931) were the first to compare articulatory defective children and normal speaking children in terms of their speech discrimination performance. They found that the normal speakers performed better than the defective group using 366 pairs of nonsense syllables as stimuli. This trend persisted among the three age groups tested, ranging from junior primary to fifth grade. Later, Hall (1938) obtained results contradictory to those of Travis and Rasmus, finding no differences between experimental (articulatory defective) and control (normal) groups. Hall

used the Travis-Rasmus test as well as a special test devised for her study. The new test employed nonsense words couched in sentences as the discriminative elements.

Further investigations into the relationship between speech articulation skills and speech discrimination skills reported contradictory results. Some studies (e.g., Aungst and Frick, 1964; Farquhar, 1961) reported no relationship between the variables. Other investigators (e.g., Cohen and Diehl, 1963; Scheifelbusch and Lindsey, 1958) found a relationship between the abilities to articulate and to discriminate among speech sounds.

A major difference in experimental procedures may account for the differences in the results of the investigations into the relationship between speech articulation and discrimination. Results were negative in those investigations which specified the misarticulation of only one phoneme as the criterion for placement in the articulatory defective group. For example, Aungst and Frick (1964) concentrated on the ability to articulate the glide /r/. Farquhar (1961) specified the misarticulation of only one sound in two word positions as the criterion for placement in the articulatory defective group.

On the other hand, Cohen and Diehl (1963) reported a significant difference ($p \leq .05$) in performances on Templin's (1943) speech discrimination test between two groups of children. One group was normal, and the other group was comprised of children having severe functional

speech articulation problems. The criterion for placement in the articulatory defective group was five or more consonant sounds in error. Scheifelbusch and Lindsey (1958) found that their articulatory defective group achieved poorer scores on a picture discrimination test than a matched group of normal speakers. The mean number of speech articulation errors among the articulatory defective group was 11.58.

Other investigations indicated that the relationship between speech discrimination and speech articulation may be confined to certain sub-groups among those who manifest speech articulation difficulties. Prins (1963) did not find a significant difference (p=.05) between his experimental and control groups. He did, however, find a strong negative correlation between performance on the Wepman (1958) auditory discrimination test and high proportions of phonemic substitution errors which involved the single articulatory feature of place of articulation.

Woolf and Pilberg (1971) used self-made tests to determine whether a relationship existed between several discrimination tasks and /r/ production as measured by McDonald's (1964) _A Deep Test of Articulation_. The results indicated that each of Woolf and Pilberg's tests were significantly different from each other (p=.01) in terms of the performances they elicited from the experimental subjects. These tests were intended to measure external discrimination, internal discrimination and "sound evaluation."

In the "sound evaluation" test, a child evaluated his own spoken responses as presented by means of a tape recording. Woolf and Pilberg's data supported the contention that there may be a relationship between various sub-populations of individuals having articulation disorders and those having speech discrimination difficulties.

Lapko and Bankson (1975), using the Farquhar-Bankson In-Depth Test of Auditory Discrimination (Farquhar and Bankson, 1971) found a positive correlation (r=.55, p=0.01) between the ability of children to discriminate their own /s/ productions and the consistency of their /s/ misarticulations. A mild positive relationship between consistency of /s/ misarticulation and stimulability of standard /s/ production (r=.41, p=0.05) was found as well. Performance on external discrimination items was somewhat better than that for internal discrimination, but significant correlations were not found between the composite scores from the internal monitoring and external monitoring sections of the test battery used. Scores on the four external discrimination sections were compared to scores on the three internal discrimination sections of this test.

Weiner (1967), Winitz (1969), and Powers (1971) critiqued the literature relative to the relationship between speech articulation and speech discrimination. Weiner (1967) noted that speech discrimination and speech articulation were similarly age related in his review of sixteen studies (e.g., Aungst and Frick, 1964; Cohen and

Diehl, 1963; Farquhar, 1961; Prins, 1963). This positive relationship was almost invariably found in children below age nine and almost never found above. One study that Weiner reviewed (Dickson, 1962) did not evidence a relationship (p=0.05) between the variables. Dickson, however, did not specify his criteria for placing subjects in the "articulation defect" group. Further, Dickson did find a significant (p=0.01) difference in motor proficiency, as measured by the Oseretsky Tests of Motor Proficiency (Doll, 1946), between the normal speaking group and the articulatory defective group. Weiner felt that the evidence he examined supported a hypothetical link between speech discrimination and articulation defects. This relationship seemed to exist in the primary age groups. Beyond this, the studies did not indicate such a relationship. An important point developed by Weiner was that the relationship between the two variables was most meaningful where the articulation defect was sizeable. The tendency among the studies reviewed was for a positive relationship to result from the use of children with the larger number of sounds in error. Thus, the child's age and the number of sounds misarticulated appeared to be equally important parameters. For example, Kronvall and Diehl (1954) inferred a relationship between the variables when they defined "articulatory defective" as meaning "four or more sound errors." On the other hand, studies which found no positive

relationship (e.g., Aungst and Frick, 1964) used a criterion of only one or two sounds in error.

Winitz (1969) reviewed seventeen studies (e.g., Cohen and Diehl, 1963; Schiefelbusch and Lindsey, 1958; Templin, 1957; Travis and Rasmus, 1931) and noted that most articulatory defective groups have been studied without regard to the specific sounds in error. Winitz posited that the sound discrimination difficulties of children with articulatory impairments were results of the articulatory errors themselves. He felt that speech sound discrimination might be more difficult following exposure to conditions which permit the correct learning of a sound. He investigated this hypothesis in two studies (Winitz and Bellrose, 1962, 1963), the results of which fell short of significance. One of the studies reviewed by Winitz was Aungst and Frick's (1964) work. Although, as mentioned above, they concentrated on one phoneme, /r/, their findings are of interest relative to Winitz's hypothesis. They presented 27 clinical clients, ranging in age from eight to ten years, with four tasks: (1) a comparison of the subjects' production to that of the experimenter; (2) a delayed judgment of the accuracy of the subjects' responses, using a tape recording, by the subjects themselves; (3) an instantaneous judgment by the subjects of their own responses; and (4) the Templin (1957) auditory discrimination test. The conclusions of Aungst and Frick were that the ability to discriminate between paired auditory speech

stimuli presented by another speaker was unrelated to the ability to judge one's own speech productions as correct or incorrect. Also, the abilities measured by the Templin test were unrelated to consistency of /r/ articulation. Further, the ability to judge one's own speech production was significantly ($p=0.01$) related to consistency of /r/ articulation.

Aungst and Frick's data suggested that the learning of an inaccurate articulatory posture when attempting to approximate the adult target sound in a given context may affect the accuracy of subsequent discrimination of that sound in similar contexts. Thus, the articulatory error theoretically contributed to the discrimination error.

Powers (1971) reviewed research findings to date (e.g., Cohen and Diehl, 1963; Farquhar, 1961; Prins, 1963; Templin, 1957; Weiner, 1967) and found that results were conflicting and inconclusive. She felt that the evidence weighed more on the side of an inferiority in speech sound discrimination in functional articulation defectives as compared with normal speakers. It was her opinion that the existence of a generalized inferiority in the discrimination of speech sounds among individuals having articulatory disorders was doubtful. Like Weiner, however, she felt that such a defect probably existed among the more severe articulation cases. Powers went on to speculate that the functional articulation defective person may have speech sound discrimination problems specific to that sound or

group of sounds which is misarticulated. While Winitz (1969) speculated that speech sound discrimination in children with articulatory problems resulted from the articulatory errors themselves, Powers posited that an individual may misarticulate certain speech sounds because of poor discrimination. She felt that the data pointed to limited and selective speech sound discrimination difficulties among functional articulation defectives.

Lewis and Kelly (1974) found a significant relationship ($p=0.05$) between the performances of a group of adult subjects on the noise subtest of the GFW and on the oral stereognosis test of Ringel, House, Burke, Dolinsky, and Scott (1970). The relationship between oral somesthesis and speech discrimination serves as an indirect means of describing a relationship between speech discrimination and articulation, since Ringel _et al._ (1970) found that articulatory defective subjects made more oral stereognosis errors than did subjects with normal speech.

In conclusion, research on the relationship between speech articulation and speech discrimination indicates that the two behaviors are similarly age related up to about nine years of age. Subjects who have functional articulation disorders may well demonstrate impaired speech discrimination abilities upon being tested. This has been reported to be a particularly likely occurrence in those cases in which the functional disorder of speech articulation involves the substitution of three or more phonemes.

Studies in which the subjects had only one or two error sounds yielded equivocal results.

Speech discrimination and language. Several studies have been oriented to determining the existence of a relationship between speech discrimination and other variables associated with communication, such as intelligance, psycholinguistic abilities, dialect and, as a related variable, environment.

Schlanger and Galanowsky (1966) compared the performances of mentally retarded children (N=85) to the performances of normal children (N=85) in response to four tests of speech sound discrimination and one test of articulation. The results supported a positive relationship (p=0.01) between mental age and speech sound discrimination in both groups. The mentally retarded children, whose IQ's averaged 60 (S.D.=15.6), gave poorer performance on the discrimination tests than did the normal group, whose IQ's averaged 104.4 (S.D.=16.3). There was no significant relationship inferred between mental age and speech articulation in the normal group. Speech articulation was significantly poorer (p=0.01) for the retarded group as compared to the normal group.

Rechner and Wilson (1967) administered the Illinois Test of Psycholinguistic Abilities (Kirk, McCarthy and Kirk, 1968), or ITPA, to forty first grade boys and girls, ages six and one-half to seven and one-half years. They

compared the performances on the ITPA with the results of speech articulation and discrimination testing. The results suggested that, regardless of speech articulation skills, children with adequate discrimination performed better on the ITPA than did those with deviant discrimination.

In a related study, Perozzi and Kunze (1971) compared the performances of thirty kindergarted students on the ITPA and and two speech sound discrimination tests. One of the discrimination tests was an abbreviated form of Templin's (1957) test. The other discrimination test was synthesized from Templin's test, consisting of thirty-three nonsense syllables with no linguistic meaning. Significant ($p=0.05$) correlations were found between the expressive language portions of the ITPA and the speech discrimination tests while similar results were found for the receptive language subtests of the ITPA.

The study of auditory figure-ground ability is especially pertinent to the present investigation since an independent variable is competing noise. Research suggests that the ability to sort acoustic stimuli into the foreground or background is a critical variable in the development of the ability to communicate (Myklebust, 1968; Sabatino, 1968).

Goldstein (1948) represented the Gestalt view of sensory performance. In describing the effects of focal brain damage on performance, he noted that all damage in

the nervous system, especially in the cerebral cortex, disturbs the normal relation of figure and ground process.

Broadbent (1958; 1962) presented a general theory of attention, memory, and filtering. The filtering mechanism was said to be necessary to protect a sensory channel against overloading. Filtering occurs before input to memory is accomplished, according to Broadbent. There remains some theoretical dispute about the temporal aspects of the occurrence of signal filtering relative to the analysis and storage process. Studies of how the brain matures to develop figure-ground organization are few (March, 1973).

Gray, Michael and Sklar (1968) investigated the effects of white noise on the speech reception thresholds of forty-eight normal and brain-injured children. The subjects, ranging in age from five to eight years, were divided into four groups: brain-injured with normal hearing sensitivity, brain-injured with impaired hearing sensitivity, the non-brain-injured with normal hearing sensitivity, and non-brain-injured with imparied hearing sensitivity. The results suggested that when brain-injured children were compared to non-brain-injured children, given equivalent hearing sensitivity, no significant differences were found in their speech reception thresholds in either quiet or noise.

Sabatino (1968) constructed a test of auditory perception, in which CVC nonsense syllables and monosyllabic

words were presented to experimental subjects. He administered his test under both quiet and noise conditions to matched groups of thirty normal children and thirty children diagnosed by psychometric testing as having minimal brain dysfunction. From the results, Sabatino concluded that the test of auditory perception in noise was the best discriminator between the groups. With the auditory test, twenty-nine of the children were identified, whereas the *Bender Visual Motor Gestalt Test* (Bender, 1946) identified only five.

Normative data supplied with the *Goldman-Fristoe-Woodcock Test of Auditory Discrimination* show the same age-related acceleration in performance for both the quiet and noise subtests. As described earlier, the greatest change took place among children between four and twelve years of age.

Marsh (1973) constructed a test of auditory figure-ground perception (AFG) using bisyllabic words as stimuli. One of Marsh's stated goals was to design a suitable measure of AFG perception which avoids use of other modalities. She presented spondee words mixed with white noise dubbed on a tape recording at signal-to-noise ratios of 2 dB, 6 dB, and 11 dB to 210 subjects from kindergarten through the third grade at an elementary school. The results indicated a negative correlation ($r=-0.47$; $p=0.01$) between errors on the AFG test and chronological age. Comparing

the AFG test to the GFW, Marsh commented that her test was easier because it required a verbal response instead of a visual-motor one. These data suggest a developmental trend in auditory figure-ground ability.

Willeford (1976) reported some clinical findings regarding the performance of seven learning disabled children, ranging in age from five to thirteen years, on four speech perception tests. The tests included dichotically competing sentences, filtered speech, binaural spectrum fusion and altering speech as stimuli. All of the test cases had normal non-distorted speech discrimination scores. The results indicated highly variable performances among the children on the various tests of speech perception. Some children achieved perfect scores on several of the tests and very poor scores on one. All but one of the children also achieved very low scores on ITPA auditory subtests.

A test of central auditory abilities was developed by Costello and Flowers (1970) for use with young children. The purpose of the test was to identify those children who have unusual auditory difficulties which cannot be explained on the basis of a peripheral hearing loss, an intellectual or a psychological deficit. The test consisted of competing message and low-pass filtered speech tasks. Costello (1977) reported that the test was useful in identifying children with language or learning problems. The picture-pointing

test format made it suitable for use with children of kindergarten age.

Black dialect has been shown to vary in grammar and phonology from General American English (Pederson, 1964; Labov, Cohen, Robins, 1965). Berlin and Dill (1967) found that special instructions and feedback improved the performances on a second trial of the <u>Wepman Auditory Discrimination Test</u> (Wepman, 1958) by an experimental group of twelve Black children. A group of ten Caucasian children received the same treatment but showed no similar improvement. The speech sound discrimination test was administered via live voice, according to instructions. Berlin and Dill concluded that an inattentive child should receive admonishing instructions and verbal reinforcement before a second trial on the Wepman test is presented. Goettesman (1972) used a specially constructed speech discrimination test to compare performances of three groups of children (N=120). The two goups of Blacks included forty subjects who spoke standard English and forty subjects who spoke Black English. Forty White subjects spoke standard English. The children responded to discrimination items presented by two different adult speakers, one who spoke Black English and one who spoke standard English. The stimulus presentation intensity levels were not specified. The results of the Goettesman study strongly suggested that Black English speaking children perform better in response to words which

contrast in their own dialects. That group performed significantly ($p \leq 0.05$) lower than the other two groups on words pronounced as homonyms in Black English dialect but not in General American English. Neither group of Black children showed a significant ($p \leq 0.05$) difference in performance in response to the two types of speakers. White children performed best in response to stimuli delivered in standard English ($p \leq 0.05$).

Factors influencing speech discrimination abilities have been related to an individual's environment. Holm and Kunze (1969) compared abilities of a control group of normal hearing children with no history of middle-ear disease to an experimental group of children with a history of fluctuating conductive hearing losses caused by recurring otitis media. The groups were matched with respect to age, sex and socioeconomic status. The results implied that, while there was no significant ($p \leq 0.05$) difference with respect to visual and motor skills, the experimental group performed poorer than the control group on measures of speech, language and listening skills. The long-term effects of this pathological factor are not clear in view of the results of this study. It appears, however, that environmental health factors can have an affect on speech discrimination skills.

Environmental noise levels of sufficient intensity can do more than mask speech (Miller, 1974). It may be hypothesized that noise levels sufficient to interfere

with auditory communication by adults should be sufficient to interfere with the acquisition of such communication skills in children. In addition, Miller (1974) posited that ambient noise levels which require a speaker to shout (e.g., 75 dB A) probably force the speaker to stop speaking, to change the content of the conversation or to speak only when necessary.

Wachs, Uzgiris and Hunt (1971) conducted a cross-sectional, correlational study of 102 infants from seven to twenty-two months of age. They found that noise level in the home was a strong environmental influence. This study was limited by the use of questionnaires and subjective ratings of noise levels. The authors concluded that high ambient noise levels in the home had negative effects on psychological development, including the development of speech and language as well as attentional control.

The effects of noise levels in the home on speech discrimination were studied by Cohen, Glass and Singer (1973). The subjects were fifty-four elementary school children who were in grades two through five. They all lived in four thirty-two floor apartment buildings located adjacent to a heavily travelled downtown expressway. A-weighted sound pressure levels, measured at hallway windows (windows closed), were inversely related to floor number ($r=-.90$). The sound pressure level on the eighth floor, for example, was 66 dB (A). Conversely, on the thirty-second floor, the sound level was 55 dB (A). The authors

administered the Wepman Test of Auditory Discrimination to children who lived in these conditions. A significant correlation ($r=0.48$; $p < 0.01$) was found for children who had lived in the apartment four years or longer. Cohen et al. subjected their data to additional analysis in order to identify additional factors, such as parents' education or carbon monoxide level; and only one, noise level/floor level, emerged as the most significant variable. Although their results appear convincing, interpretation of these data is mitigated by the fact that the audiological status of the subjects was not carefully determined. This leaves open the possibility that the subjects, especially long-term residents, had noise induced hearing losses. Further, the type of judgment demanded by the Wepman test may have been too difficult for some of the subjects.

Deutch (1964) opined that children raised in a noisy environment may have problems in auditory attentiveness. These problems could be related to repeated attempts to cope with unwanted sounds, leading to a failure to attend to all sounds. The results may lead to a reduction in speech discrimination performance. One could argue that, conversely, children raised in noisy environments might acclimate to the noise and thus become better discriminators.

The literature reviewed here supports a relationship between speech discrimination and speech articulation in children up to age nine, if the articulation defect is

sizeable. There continues to be some question about whether a speech discrimination defect is causally related to poor speech articulation or whether the inverse is true. Comparisons of subgroups within both the speech discrimination population as well as the articulation population may serve to clarify the relationship between the two behaviors. In addition to speech articulation, other variables have been shown to have a relationship to speech discrimination. Intelligence, dialect, psycholinguistic abilities and environment have emerged as factors which influence skill in speech discrimination.

Factors Involved in Speech Discrimination Testing

A trend is seen in the development of speech discrimination tests in which research demonstrates the existence of confounding variables in testing and existing tests are modified to accommodate thse variables. The literature of speech discrimination testing identified azimuth of signal propagation, closed-set versus open-set format, response mode, stimulus material and speaker characteristics as variables which affect listener performance.

Azimuth. Research on the nature of tests of speech discrimination has shown that signifcant differences in performance can be obtained by manipulating variables

related to the physical characteristics of the test. Norlund and Fritzell (1963) found that azimuth, or angle of incidence of signal propagation towards the head, has a substantial influence on speech signals. They used an artificial head fitted with condenser microphones for "ears." Findings indicated that intelligibility of phonetically balanced words was better with binaural listening at all azimuths tested. Information reaching the ear turned toward the sound source was better, as verified spectrographically, than information reaching the ear turned away from the sound source. Changes were most conspicuous at frequencies greater than 2KHz. The inter-ear difference did not occur at 0° and 180° azimuths.

Closed-set versus open-set format. The closed-set design appears to be preferable to the open-set design in speech discrimination testing. Schultz and Schubert (1967) saw five advantages of the closed-set over the open-set design. They felt that the closed-set design (1) minimized "score inflation" caused by word frequency errors; (2) afforded control of item difficulty; (3) afforded single sound substitutions, systemizing item analysis; (4) focused attention on the relationship between test goals and the criteria for selection of response foils; and (5) allowed automation of procedures for presentation, scoring and evaluation.

Word frequency effect is the extent to which word discrimination scores are affected by the listener's familiarity with the words used as stimuli. Pollack, Rubenstein and Decker (1959) have shown that word frequency effects are absent or minimal with known (closed) message sets, but that strong word frequency effects may be observed with unknown (open) sets. These data are pertinent to the present study, since the speech discrimination tests used here are of the closed-set design. According to Pollack et al. (1959), in a closed message set design, the prime factor determining intelligibility is phonemic interconfusability among words. In open message sets, word frequency effect is a highly influential variable. Oyer and Doudna (1960) studied the effects of word familiarity on the responses made to the C.I.D. Auditory Test W-22 (Hirsch et al., 1952). Using hearing impaired subjects, they reported that words which occurred more frequently in English elicited a higher number of word omission or no response errors from their subjects. Word substitution responses tended to increase as frequency of word occurrence decreased. Egan (1948), Owens (1961), and White (1974) have also found that word familiarity is an important factor in speech discrimination test performance. Jones and Studebaker (1974) favored the closed-set for use with severely hearing impaired subjects whose level of performance was low (but not 0%) on open-set speech discrimination tests. Further, they found that closed response test

scores were positively correlated with other data dependent on hearing function, whereas the open response scores were not. Jones and Studebaker's subjects determined their own listening intensity levels. Responses required marking a printed word as correct. Miller, Heise and Lichten (1951) determined that the range of alternative response possibilities had an effect on performance by their sample; and Ross and Lerman (1971) have stated that having a closed-set of response alternatives can result in better performance than having an open-set.

Fairbanks (1958) described a method of testing a restricted aspect of speech reception which he referred to as phonemic differentiation. The *Rhyme Test* was developed in response to a need for experimental materials in which (1) the stimulus was the spoken word; (2) the response was recognition of the word; (3) the response was dependent upon the initial consonant and consonant vowel transition; and (4) the subjects' task would bear valid relation to the discrimination demands of real speech. Fairbanks also endeavored to make auditory-phonemic factors weigh heavily in the score and make linguistic factors of higher order weigh lightly. The *Rhyme Test* is of the completion type. Subjects complete the spelling of a word by adding the initial consonant following its auditory presentation within the context of the word. Great care was taken to avoid spelling confusions. Stimulus words were chosen with

respect to their frequency of occurrence. House *et al*. (1965) modified the *Rhyme Test* in an effort to develop a testing procedure that could be used routinely by relatively naive talkers and listeners to evaluate the performance level of speech communication systems. Their test differed from Fairbanks' (1958) in terms of the constraints imposed on the words and word lists. The words were chosen with consistency of orthographic representation of the vowel nuclei in mind. The word lists were not phonetically balanced but represented sounds from the "major categories" of speech sounds. The nature of subjects' task was also different from that of the *Rhyme Test* in that a closed-set paradigm was used. House *et al*. speculated that their word lists could be administered repeatedly (e.g., at various signal-to-noise ratios) without improvement in scores related to practice effects. Their data showed that the magnitude of the standard error decreased as the number of exposures to the test materials increased and as the number of listeners tested increased. However, the differences between scores on repeated trials were not significant ($p=0.01$). House et al. (1965) tested discrimination of both initial and final consonant sounds among their experimental subjects, whereas Fairbanks (1958) had only investigated discrimination of initial consonants and consonant-vowel transitions.

Response mode. The type of response required by a test has been shown to be a factor contributing to performance variability (Aungst and Frick, 1964; Woolf and Pilberg, 1971). Weiner (1967) felt that the picture pointing type test, especially when used with children, was of questionable validity because of its apparent near identity with vocabulary knowledge. However, Ross and Lerman (1971) described speech articulation defects and limited graphic abilities as factors which tended to preclude oral or written responses as accurate indices of speech discrimination. Schlanger and Galanowsky (1966), as mentioned earlier, found a relationship between mental age and speech discrimination. They used the picture pointing type test. Templin (1957) and Mills (1975) have both criticized the Travis and Rasmus (1931) type test on the grounds that the task of attaching a label of "same" or "different" to a pair of nonsense syllables is too difficult for young children. Templin's speech discrimination tests use both the "same/different" judgement response to nonsense syllables (1943) and the picture pointing response (1957). The "same/different" response to nonsense syllables is also used in Wepman's (1958) speech discrimination test. The Boston University Speech Sound Discrimination Picture Test (Pronovost and Dumbleton, 1953) combines the "same/different" judgement with a picture pointing response. The pictures used are silhouettes arranged in pairs, the parts of which are identical (same) or minimal pairs (different).

Locke (1980a) stressed the importance of testing children's perception of speech as a part of the clinical evaluation process. Although he stated that the picture identification test format was more desirable than the "same/different" (AX) format, Locke discussed some variations on the AX format in a related article (Locke, 1980b). These variations were said to help specify the direction of the misperception of specific speech sounds. Two of these alternative formats. the ABX and the oddity task, involved the presentation of triplets of syllables to a listener. The listener determined which two of the three were most similar. A third variation, the four-interval AX test, involved the presentation of two pairs of syllables. The listener identified which of the two pairs was most nearly alike. Locke's (1980a) preliminary data obtained with fourteen pre-school and school-aged children indicated tentatively that there was little variation across the several tasks. Locke's attempt to study phonological substrates of speech development in children, while a step in the right direction, suffered from some basic problems. First, the several test format variations were to be subject to the same limitations regarding a child's understanding of task requirements of the AX design. Second, Locke apparently mixed meaningful and non-meaningful stimuli (e.g., pig-zig) in his stimulus pairs. Finally, all stimuli were administered live-voice, face-to-face with no control of stimulus intensity or visual cues.

Stimulus material. The first formal test of speech discrimination was an attempt to understand more fully the relationship between articulation and discrimination of speech sounds. The test, designed by Travis and Rasmus (1931), was a comparison of speech sounds in 336 pairs of nonsense syllables. It was shortened to seventy items by Templin (1943).

Black (1952) examined the relationship between the intelligibility of words and the aspects of syllabic pattern, word familiarity and phonetic characteristics. His results, based on the responses of 80 panels of naval trainees, indicated that words having many sounds and two syllables were more intelligible than words with few sounds and only one syllable. Black also reported that more familiar words were more accurately identified. Finally, Black identified certain phonetic elements that were associated with enhanced intelligibility and others that were associated with detracted intelligibility in his investigation.

Templin developed a speech discrimination test for children. She shortened her seventy item test (Templin, 1943) to fifty items and used it with children aged six to nine years (Templin, 1957). Hall (1938) and Mase (1946) developed tests in which the stimulus words were contained within a sentence; and Pronovost and Dumbleton (1953), Templin (1957) and Scheifelbusch and Lindsey (1958) presented picture speech discrimination tests which have been

models for tests presented more recently by Seigenthaler and Haspiel (1966), Ross and Lerman (1971) and Goldman, Fristoe and Woodcock (1971).

Relative to stimuli, Sabatino (1969) concluded that using consonant-vowel-consonant (CVC) syllables is a poor way to assess auditory perception, since the task requires no interpretation of the stimuli. In Sabatino's study, however, stimuli were presented at 60 dB HTL (A.N.S.I. 1969) without respect to the individual subject's speech reception threshold. Templin (1943) used CVC syllables specifically because they were devoid of linguistic meaning, as did Perozzi and Kunze (1971).

Lehiste and Peterson (1959) felt it was important to point out that "phonetically balanced" nonsense syllables represented a contradiction in terms, since phonetics refer to psychological and acoustical phenomena which, by nature, cannot be balanced by linguistic occurrence. They preferred the terms "phonemic balance," arranging their consonants around vowel nuclei. These syllables were denoted as consonant-nucleus-consonant (CNC) syllables.

A novel approach to stimulus development was presented by Speaks and Jerger (1965). Rejecting both the CVC (or CNC) and single word designs, Speaks and Jerger's message set stimuli contained synthetic sentences. These sentences were constructed as approximations to real sentences solely on the basis of conditonal probabilities of

word sequences. Their test required a subject to identify among ten responses available in written form which of several alternative messages were presented auditorily. Thirty subjects demonstrated that as the amount of information in artificial sentences decreases, performance in identification of these sentences improves.

Speaker characteristics. Kruel, Bell and Nixon (1969) and Rintelmann, Schumaier and Jetty (1975) demonstrated that a change in the speaker who presents the stimuli can result in a change in the subjects' performances; and, in the Kruel, Bell and Nixon (1969) study, a change in the carrier phrase coincided with a change in subject performance. Kruel *et al*. used recordings of the *Modified Rhyme Test* (Fairbanks, 1958) for speech stimuli, while Rintelmann *et al*. presented recordings of the *N.U. Auditory Test No 6* (Tillman, Carhart and Wilber, 1966). In a study of variables affecting the performance of subjects on speech discrimination tests, Schwartz and Goldstein (1974) presented the GFW to seventy-two pre-schoolers. Results suggested that the context of stimulus presentation affected accuracy of performance. All of the subjects made significantly more ($p \leq 0.01$) errors in the context using limited grammatical and phonetic cues. Stimulus contexts included the paired comparison task, carrier phrase context and sentence context. Picture pointing responses were required in all contexts.

Goldman, Fristoe and Woodcock (1971), in an effort to develop a discrimination test relatively unconfounded by the factors mentioned above, developed their Test of Auditory Discrimination. It consists of a series of plates, each one containing four black and white line drawings. The child looks at a plate containing four pictures and is asked to point to the one that matches the word he hears. The test was standardized on 506 subjects from four to twelve years of age, as well as on a group of adults. According to the authors, extraneous variables are minimized in three ways. First, with each subject, the adequacy of the test materials is evaluated and improved if necessary. A training session insures that the subject is familiar with the stimulus words and their corresponding drawings. Second, administration is controlled through use of a pre-recorded tape, so that the voice and carrier phrase are standardized. Finally, the authors claim that a different aspect of memory is added (the picture pointing response), avoiding what the authors call a "rather artificial" type of auditory memory task found in many tests (Goldman, Fristoe and Woodcock, 1971). The range of response alternatives is limited to four per item.

Similar in construction, the Word Intelligibility by Picture Identification (WIPI) test was developed by Ross and Lerman (1971) as a means of addressing problems specifically related to the testing of auditory discrimination

in children. They also used a picture pointing format but with a range of six possible responses per plate. They considered this an improvement over conventional measures for three reasons: vocabulary limitations related to language retardation are minimized; oral responses from children with long standing hearing losses and concomitant articulatory problems are eliminated; and written responses are not necessary. The WIPI pictures are drawn in color, in contrast to the black and white GFW pictures. In interpreting the use of their own instrument, Ross and Lerman stated in the WIPI test manual that discrimination scores obtained with this test cannot be considered equivalent to the scores obtained with conventional speech discrimination lists. Since the WIPI test is a closed-set, chance scores are approximately 18%.

The WIPI was not standardized as was the GFW. Instead, the WIPI was evaluated on sixty-one hearing impaired subjects ranging in age from four years, seven months to thirteen years, nine months. Stimuli were presented at 40 dB re: SRT. Pearson product-moment correlation coefficients of the four lists indicate the four WIPI lists to be highly equivalent ($.84 \leq r \leq .95$).

The authors of both the WIPI and the GFW state in the manuals that they recognize the problems of conducting concurrent validity studies with their instruments. The GFW authors felt that correlation of their test with existing measures of speech discrimination would be difficult to

interpret, since other test results would have been more confounded than those of the GFW. They addressed this problem by using the judgments of expert clinicians at the Bill Wilkerson Hearing and Speech Center in Nashville, Tennessee. A point-biserial correlation of .68 was obtained between clinical judgment and t-scores on the Quiet Subtest (Goldman, Fristoe and Woodcock, 1971). The WIPI authors made no attempt to study the validity of their instrument by comparison with some other measure of speech sound discrimination, leaving that for future research. Interestingly, both tests were published in 1971.

The Performance-by-Intensity Function

Speech discrimination testing has become generally regarded as an essential item in the diagnostic audiological battery (Tillman and Olsen, 1973). Silverman and Hirsch (1955) felt that the importance of speech stimuli was derived from its nature as a particular kind of auditory stimulus, rather than its more pragmatic acoustic properties. They felt that diagnostic speech discrimination testing became more functional as the stimuli used became more similar to everyday conversational speech. As regularly applied, discrimination testing is presented to the client at a level where maximal performance is usually observed.

The social adequacy index for hearing (SAI) was presented by Davis (1948) in an effort to quantify an individual's degree of handicap in various social situations. Davis felt that consideration of the dynamics of everyday speech was important in the evaluation of hearing. The SAI is a measure that is computed from the speech reception threshold and from speech discrimination testing at three intensity levels relative to the patient's SRT. These levels were 33 dB, 48 dB and 63 dB in the original research (Davis, 1948). The SAI was conceived as a measure of the average probability of hearing a word accurately in all types of everyday listening conditions. A table was devised to simplify testing. Difficulties developed with its use because of a limited sample of hearing impaired subjects used for reference. Further, the use of old recordings of the Harvard PB-50 (Egan, 1948) word lists to obtain the data on which the degrees of handicap were based has been questioned, since they were not standardized well enough to measure discrimination adequately (Davis, 1970).

Investigations have suggested that changes in listener performances accompanying successive increases in speech stimulus presentation level can be of diagnostic significance. The change in performance has also been called the performance-by-intensity (PI) function. Jerger and Jerger (1971) reported differences in PI functions

related to differences in the types of hearing loss suffered by their subjects. PI functions for phonetically balanced word lists (PI - PB functions) obtained with patients having lesions of the VIIIth cranial nerve showed a "rollover" effect as intensities were increased to very high magnitudes (100 dB HTL, ANSI 1969, or more). All of ten subjects having VIIIth nerve lesions, as well as all of six patients having bulbar lesions, demonstrated a rollover effect. The sites of the subjects' lesions were confirmed by neurological evaluation. Jerger and Jerger included one six-year-old child among their subjects and modified the discrimination materials accordingly. They anticipated that 70% to 80% of patients with retrocochlear auditory disorders would show the rollover effect. In patients having cochlear disorders, further increases in speech stimulus levels did not result in significant reductions in speech discrimination.

Dirks (1978) summarized the general configurations of PI-PB functions. For conductive hearing impairments, the configuration is generally of the same shape as for normal hearing, the difference being that equivalent percentages of correct stimulus identification are obtained at higher stimulus intensity levels when the hearing impairment is conductive. A cochlear lesion has the effect of lowering the maximum discrimination score. A slight degree of rollover may be seen. A retrocochlear

lesion results in a lower maximum discrimination score. A pronounced rollover effect usually occurs, often resulting in near zero speech discrimination scores at intensities around 100 dB HTL (ANSI 1969).

An attempt to explain the rollover phenomenon was put forth by Martin and Pickett (1970). They hypothesized that for patients with sensorineural hearing losses, weak consonants produced at relatively low sound pressure levels were difficult to hear, resulting in low discrimination scores. When the intensity of stimulus presentation became great enough to allow the hearing impaired person to distinguish the weak consonants, other stronger phonemes served to mask the weak ones. This explanation would appear to be most applicable to recruiting ears, however, offering an argument for a cochler source of the rollover effect.

Estimation of the speech reception threshold (SRT) by means of examining the PI-PB function was studied by Jerger and Jerger (1976). They felt that describing the PI-PB function should be standard procedure in audiometric testing. Deriving an SRT (which was labelled PBT) from the PI-PB function, was promoted as an enhancement of the PI-PB function's clinical usefulness. Examining the data of 2,117 subjects, Jerger and Jerger obtained the PBT on all except those whose maximum discrimination score was less than 31%. Most patients showed similar thresholds for

pure-tones and PB words. Most patients who had large differences between thresholds for pure-tones and PB words demonstrated a greater loss for PB words than for pure-tones. The PBT was considered in agreement with pure-tone average if the difference between the two did not exceed 10 dB. The results of this study were in general agreement with Carhart's (1946) description of the relationship between thresholds for speech and those for pure-tones. Jerger and Jerger felt that the benefits derived from determining PBT included provision of a cross-check on pure-tone threshold level and supplementation of information about the audiometric contour. Jerger and Jerger promoted the routine use of PI-PB testing. They stated that information provided by PI-PB testing included speech reception threshold, maximum speech intelligibility score and the presence or absence of rollover.

PI-PB functions were also used as a means of comparison for word lists (Rintelmann, Schumaier and Jetty, 1974; Speaks, 1967b). Sanderson-Leepa and Rintelmann (1976) studied the performances of normal hearing children in response to three speech discrimination tests. PI-PB functions were graphed as one means of comparing the word lists. Sanderson-Leepa and Rintelmann felt that it was important to generate the PI functions to establish how children ranging from pre-school through school-age respond to various types of speech discrimination tests. Sanderson-Leepa and Rintelmann compared the WIPI, the <u>Phonetically</u>

Balanced Kindergarten Word Lists (PBK-50) (Haskins, 1949) and the Northwestern University Auditory Test No. 6 (NU-6). They reported that the WIPI appeared to yield higher and less variable mean discrimination scores for children aged three and one-half years. The WIPI and the PBK-50 appeared to be appropriate for use with children aged five and one-half, provided the children have good speech and normal language development. Results with the three remaining groups of children, aged up to eleven and one-half years, indicated that the WIPI or the PBK-50 were the preferable instruments, since they were associated with higher mean discrimination socres at all sensation levels than the NU-6.

Rintelmann, Schumaier and Jetty employed PI-PB functions as a means of determining the test-retest reliability of the N.U. Auditory Test No. 6. Speaks (1976b) used PI-PB data to compare listeners' performances with lists of synthetic sentences (Speaks and Jerger, 1965), spondaic words and phonetically balanced monosyllabic words (CID Auditory Tests W-1 and W-22, Hirsch, Davis, Silverman, Reynolds, Eldert and Benson, 1942). The synthetic sentences were associated with higher percentage scores at any presentation level in Speak's study.

Performances of ten college students were compared in a pilot study for the present investigation (Culbertson and White, 1980). The students were presented with the GFW and the WIPI in order to generate performance-by-

intensity data. The subjects were members of the student body at Michigan State University, East Lansing, Michigan. They ranged in age from twenty to twenty-seven years, with a mean age of 22.1 years. Each subject was given a pure-tone audiometric screening before presentation of the stimulus materials. Screening level was 20 dB HTL (ANSI 1969).

The subjects were seated in a double-walled sound suite and presented the stimuli in their right ears only through a Telephonics TDH-39 earphone mounted in an MX-41/AR cushion.

The performance-by-intensity curves generated in the pilot study showed that for these subjects the mean scores compared quite closely at each intensity level tested. The linear portions of the curves appeared between 0 and 8 dB HTL. There was only a slight difference between the slopes of these lines (6.23%/dB versus 51.5%/dB). Variability of scores was highest at either end of the straight portions of the curves and lowest at the extremes of the intensity range.

Wilson and Antablin (1980) used PI functions to compare a multiple choice-type picture identification speech discrimination list of their own construction with the *N.U. Auditory Test No. 6* (Tillman *et al.*, 1966) among a sample of verbal and non-verbal adult subjects. Adults who were verbal performed equivalently in response to the two lists. Non-verbal adults demonstrated differential

performances depending upon the response paradigm at low sensation levels. The best performances were elicited in the closed-set response to written words. At higher intensity levels, 26 dB S.L. and above, the performances in all three response paradigms were equivalent.

Thornton and Raffin (1980) constructed a table of probability values for comparing percentage scores obtained with open-set design speech discrimination lists. The goal of their effort was to provide a basis for estimating the significance between speech discrimination scores obtained with different lists. In an earlier article, Thornton and Raffin (1978) noted that a client's observed score of 100% in response to a given speech discrimination list corresponded to a "true score" (p=0.95) range of 86 to 100%. The range of probable scores for a client scoring 48% was from 28 to 69%, according to Thornton and Raffin.

Research findings relative to performance-by-intensity function testing indicate that its incorporation into the clinical diagnostic audiological test battery can provide valuable site-of-lesion information. Its limitations as a diagnostic method are that it is fatiguing to even the most mature subjects. In addition to its clinical use, the performance-by-intensity function is becoming accepted as a means of comparing speech stimuli materials used in speech and language pathology and audiology.

Summary

Auditory discrimination testing, as a term, can include probing a subject's abilities in distinguishing between all forms of auditory stimuli. Speech discrimination refers to the differential perception of speech sounds or of acoustic stimuli whose spectral characteristics approximate real speech. Word discrimination refers to distinguishing among linguistically meaningful speech stimuli. Speech discrimination has been shown by research findings to be a developmental phenomenon, beginning in earliest infancy then improving rapidly to eleven or twelve years. Further improvement with age is perhaps less remarkable. The appearance of auditory problems associated with old age brings with it various degrees of decrement in speech discrimination ability.

A relationship has been documented between speech discrimination abilities and other variables associated with speech and language. Among these variables are included speech articulation, intelligence and psycholinguistic skills, such as auditory figure-ground perception. Dialect and environment appear to be important factors as well.

The development of speech discrimination testing has been traced from its inception in the early 1930's to its present state, in which variations in the angle of

incidence between the listener and the speaker, task requirements, speaker characteristics, and carrier phrase have been identified as affecting performance. Two multiple choice-type picture tests of speech discrimination, the GFW and the WIPI, have been described in detail, since they play a particularly important role in the present study.

Finally, the relationship of speech discrimination test performance to stimulus intensity, known as the performance-by-intensity function, has been shown to be a useful clinical phenomenon as well as a means of comparing the various word lists and tests used in speech discrimination testing.

Chapter 3

METHODS AND PROCEDURES

The methods and procedures for examining the performance-by-intensity of normal hearing children in response to two multiple-choice type picture speech discrimination tests are presented in this chapter. Subjects are described first followed by a description of the stimulus materials and apparatus used. Following this, the general procedure and data collection details are explained.

Subjects

The subjects for this study were eighty elementary school students, forty of whom were equally divided into two age groups: five years (N = 20) and ten years (N = 20). Ten randomly chosen subjects in each age group were assigned to the quiet listening condition and the ten remaining subjects in each age group were assigned to the noise listening condition. A number table for random assignment (Wood, 1974) was used to assign subjects to the two listening conditions. In addition to the subjects from which the performance-by-intensity functions were generated, forty additional subjects were chosen and assigned to the two age groups and the two listening conditions in the same manner as was the repeated presentation group. This second group of forty subjects comprised the single presentation group and was presented the word lists at a sensation level at which all of the repeated presentation subjects achieved a score of 90 percent or better. Within the five-year-old group, one subject failed to reach this 90 percent criterion in response to the WIPI at any sensation level tested. Therefore, the highest (40 dB) sensation level was used to examine the effects of repeated presentations upon the performances of five-year-olds on the WIPI.

The subjects were identified by means of flyers sent out to the Flagstaff, Arizona public schools as well as to day-care centers in the Flagstaff area. Subjects' ages in the younger age group ranged from 5-0 to 5-9 ($\overline{X}$ = 5-4) and from 10-0 to 10-10 ($\overline{X}$ = 10-4) in the older group. The subjects' parents signed parental consent forms which were approved by the human subjects committee at Michigan State University, the institutional review board at Northern Arizona University, and the administrative offices of the Flagstaff, Arizona, public schools.

To participate in this study, a child (1) passed a pure-tone hearing screening at 20 dB HTL; (2) was free from a history of chronic middle ear pathology and frequent exposure to high intensity noise; and (3) was a speaker of General American English dialect.

Stimulus Materials and Apparatus

Stimulus materials. The speech stimuli used for the discrimination testing were those same stimuli used for the pilot study of this investigation (Culbertson and White, 1980). These stimuli were recorded presentations of the GFW and WIPI. In order to eliminate the variable of speaker differences, the pre-recorded audio tape supplied with the GFW was not used in this experiment. List I of the WIPI and the Quiet Subtest list of the GFW were recorded, using the experimenter's voice, through an Electro-voice E.V. 635A microphone connected to an Ampex AG 500 open reel tape recorder. The recording was made in an Industrial Acoustics Company sound resistant booth. This master recording was made to duplicate the temporal characteristics (i.e., five second interval between stimuli) and the carrier phrase used on the original GFW tape recording.

The signal on the master tape was then dubbed onto one track of a TEAC Tascam Series model 80-8 open reel 1/2 inch eight track tape recorder. A tape of "cafeteria noise"[1] was dubbed onto another track of the 1/2 inch tape and used as the noise distraction in the final mixing. The 1/2 inch tape recorder was biased for the Ampex 456 recording tape used for this procedure. This 1/2 inch tape was then played back through a TEAC Tascam Series model five

---

[1]Auditec of St. Louis, 402 Pasedena Avenue, St. Louis, MO. 63119

mixer to be dubbed finally onto a 1/4 inch open reel tape, using an Akai 4000DS Mk. II audio recorder.

The final dubbing involved playing the 1/2 inch tape program through twice so that the quiet presentations of the GFW and the WIPI and the noise presentations of the GFW and the WIPI occurred consecutively on the 1/4 inch test tape. Thus, the quiet presentations occurred first, followed by the noise presentations. A 400 Hz calibration tone was also dubbed on the test tape. The speech stimuli and the noise distraction were mixed (for the noise presentations) before the final dubbing by adjusting the mixer so that the signal-to-noise ratio was 9 dB.

The SRT stimulus tape was prepared by recording children's spondees (Utley, 1951) through an Akai model 4000DS Mk II open reel tape recorder interfaced with a Grason-Stadler model 1701 clinical audiometer. Peak amplitude of these stimuli was monitored by observing deflection of the V.U. meter on the audiometer. This deflection was maintained at a peak of 0 on the meter. A 1000 Hz calibration tone was also placed on the tape. The voice used to establish the SRT was the same one used on the tape recording of the discrimination stimuli.

Apparatus. A Grason-Stadler model 1701 clinical audiometer, calibrated to ANSI standards (ANSI, 1969) for pure-tone and speech audiometry, was used to determine the speech reception threshold and deliver the speech discrimination stimuli. The speech stimuli for these measures were presented by means of an Akai Mk. II 1/4 inch open reel tape recorder interfaced with the audiometer. These stimuli were presented in free-field through an Intensify model 1001-A loudspeaker.

The same Grason-Stadler model 1701 clinical audiometer used to present the speech stimuli was used to perform the pure-tone screening. Pure tone stimuli were delivered through Telephonics TDH-39 earphones mounted in MX-41/AR cushions.

Procedure

Each subject passed a pure-tone air conduction hearing screening at 20 dB during the intake phase of the procedure. Following parental consent to proceed, the subjects' case histories were examined to insure that they were free from chronic otitis media. Speech articulation of the subjects was screened to insure that it was within normal limits. The speech articulation screening was accomplished by presenting the screening portion of the *Fisher-Logemann Test of Articulation Competence* (Fisher and Logemann, 1971). Receptive vocabulary was assessed with the *Peabody Picture Vocabulary Test* (Dunn, 1965). Each subject's receptive vocabulary scaled score was at least 85. The experimental phase of the procedure was then conducted.

The subjects began the experimental phase of the procedure by taking a seat in the sound treated booth. Speech reception thresholds (SRT) were established by use

of pre-recorded children's two-syllable words (spondees) (Utley, 1951). The speech stimuli were presented to the experimental subjects, whether assigned to quiet or noise listening conditions, according to a plan designed to counter-balance presentations, eliminating order effects on the final mean scores. According to this plan, one-half of the subjects in each group (N = 10) received the WIPI stimuli first at a given sensation level. Following the WIPI presentation, the GFW was presented at the same sensation level. This process was repeated at each sensation level in order of increasing intensity. The remaining ten subjects in each age group received the GFW first, followed by the WIPI, at each of the sensation levels. In this manner, half of the subjects were presented the WIPI first, followed by the GFW. The other half of the subjects received the GFW first, followed by the WIPI at each sensation level regardless of listening condition.

The discrimination stimuli were presented in either quiet or noise conditions at five sensation levels relative to each subject's SRT. The five sensation levels were 8, 16, 24, 32 and 40 dB. During the presentations with background noise, the signal-to-noise ratio was maintained at 9 dB for each sensation level.

The single presentation group (N = 40) was assigned according to age and listening condition in the same manner as the experimental group. This group received the discrimination stimuli at one sensation level only. This sensation

level was the minimum level at which all of the subjects in the repeated presentation group achieved scores of 90 percent or better in the quiet listening condition. The same sensation level was used for the single presentation group subjects in each age group assigned to the noise condition.

The speech discrimination stimuli were delivered through the Intensify model 1001-A loudspeaker approximately two meters from the subjects' ears at an azimuth of 0°. Noise distraction for the noise listening condition was mixed through the same channel as the speech stimuli. Under the noise listening condition, the speech and noise came from the same loudspeaker.

Sound pressure levels were checked before each testing session with a General Radio model 1933 sound level meter. The levels were monitored by placing the sound level meter on a stand to approximate the level of the subjects' ears. The experimenter viewed the meter from a distance of several feet in order to eliminate the possibility of an artificial reading related to standing waves reflected by his body. Speech noise, generated by the clinical audiometer, was used as a signal to calibrate the equipment.

Data Analysis

Mean discrimination scores for both age groups at the five sensation levels and two listening conditions were calculated. Pearson product-moment correlation coefficients

were computed as a means of comparing the two tests at each sensation level, within each age group (Weinberg and Schumaker, 1969).

Mean performance-by-intensity functions were developed. The PI curves were compared in terms of their 90 percent correct discrimination levels and the slopes of the lines connecting the mean discrimination score points at 0 dB HTL and 8 dB HTL.

Three repeated measures analysis of variance treatments were employed to determine the effects of the factors of age and presence of background noise as the sensation level of stimulus presentation increased in five steps. The first analysis was performed on all of the repeated presentation subjects, regardless of age group. The second two treatments were used to examine each listening condition separately. The Neuman-Keuls *a posteriori* treatment (Winer, 1971) was used to examine the significance of score differences for each listening condition. The *a posteriori* treatment made it possible to identify specific areas of difference following the emergence of significant F-values.

The t-test (Weinberg and Schumaker, 1969) was used to determine the statistical significance of differences ($p \leq .01$) between the mean scores of the single presentation group and those of the repeated presentation group at the specified sensation level. This sensation level, as previously mentioned, was that minimum level at which a 90 percent or better speech discrimination score was elicited from all

of the subjects in a particular age group in the quiet listening condition.

The independent variables in this design were: (1) age; (2) listening condition; and (3) stimulus intensity. The variable of age was manipulated over two categories by selecting subjects from two age groups, five years and ten years. The variable of listening condition was manipulated over the two categories of quiet (no-noise) and noise. Intensity, the third independent variable, was manipulated over five sensation levels, 8, 16, 24, 32 and 40 dB, for both age groups and under both listening conditions.

The dependent variable for this study was the discrimination scores obtained by the subjects in the two age groups, under the two listening conditions and at the five sensation levels.

Chapter 4

RESULTS AND DISCUSSION

Five research questions represented the primary thrust of this research project. First, will there be a difference between performance-by-intensity functions obtained in the sound field with the GFW and the WIPI? Second, will the presence of background cafeteria noise affect performance on the two tests? Third, will there be a difference in performance on the two tests associated with the ages (five years and ten years) of the two groups of children tested? Fourth, can scores obtained at a specified sensation level with the GFW predict scores obtained with the WIPI at the same sensation level? Finally, will there be a difference in listeners' performances associated with single versus multiple presentations of the stimulus materials?

This chapter presents the data associated with the performance-by-intensity functions first. Included in the first section are the findings related to differences associated with single versus multiple test presentations. Secondly, differences associated with the presence of background cafeteria noise are reviewed. Next, variations in performance associated with the two different age groups are reviewed. Finally, correlational data relating the two speech discrimination tests are presented and discussed.

Performance-by-Intensity Functions

As expected, the mean performance scores for each test tended to increase as the sensation level increased. Tables 1 and 2 present the means and standard deviations of performances by all the children tested. Table 1 shows the means and standard deviations of the percentage scores for the five-year-old group in the presence and in the absence of background cafeteria noise at each of the five sensation levels. The term "percentage score" refers to the percent of correct responses elicited during each trial. Table 2 presents the mean and standard deviation data for both speech discrimination tests and both listening conditions at each sensation level for the ten-year-old group.

The standard deviations of the scores decreased as the sensation levels increased for both tests under the quiet listening condition. The same trend appears to exist for the noise listening condition. The decrease in the standard deviation of the discrimination scores obtained in the presence of background cafeteria noise was not as marked as it was with scores obtained in the quiet listening condition. Only at certain sensation levels for either age group in the noise listening condition, the standard deviation was larger than it was at the previous (lower) sensation level, unlike that for the quiet listening condition.

Data for the single presentation groups were compared with those for the repeated presentation groups for both age

TABLE 1. Means and standard deviations of percentage scores for WIPI and GFW performances by five-year-olds in quiet and noise listening conditions.

		WIPI				GFW			
		QUIET		NOISE		QUIET		NOISE	
SENSATION LEVEL (dB)		$\bar{X}$%	SD	$\bar{X}$%	SD	$\bar{X}$%	SD	$\bar{X}$%	SD
Repeated Presentation	8	61.6	17.3	57.6	13.8	79.3	12.4	66.3	13.4
	16	88.4	10.6	82.0	8.3	91.7	8.7	79.0	14.3
	24	96.4	6.9	86.4	10.7	96.0	7.5	83.3	11.8
	32	97.6	5.1	89.6	8.0	96.0	6.6	85.0	13.3
	40	98.0	5.1	90.0	7.8	98.0	3.6	86.0	12.0
Single Presentation	40	91.2	6.2	76.4	9.9	87.0	6.8	77.4	11.7

TABLE 2. Means and standard deviations of percentage scores for WIPI and GFW performances by ten-year-olds in quiet and noise listening conditions.

SENSATION LEVEL (dB)		WIPI				GFW			
		QUIET		NOISE		QUIET		NOISE	
		$\bar{X}$%	SD	$\bar{X}$%	SD	$\bar{X}$%	SD	$\bar{X}$%	SD
Repeated Presentation	8	68.4	20.6	66.8	10.5	78.7	8.5	73.3	12.4
	16	92.8	11.9	87.6	6.7	95.0	5.5	91.3	8.2
	24	99.6	1.3	93.6	4.7	97.3	2.6	96.7	5.0
	32	100.0	0	95.2	4.1	99.7	1.0	96.7	4.7
	40	100.0	0	96.0	4.6	100.0	0	97.3	4.4
Single Presentation	24	95.6	4.8	88.0	4.2	94.0	4.1	87.7	5.0

groups in order to assess the practice effect. The 40 dB SL data were chosen for the five-year-old groups while the 24 dB SL data were used for the ten-year-old groups. These levels were selected since it was at this point that 90 percent correct or better scores were elicited for both tests. Tables 1 and 2 include data for the children receiving a single presentation of each test according to their age group for comparison purposes.

Single presentation group means were significantly different ($p \leq .01$) from repeated presentation group means for the GFW in quiet and the WIPI in noise among the five-year-olds. Significant differences ($p \leq .01$) between single presentation group means and repeated presentation group means were obtained for the WIPI and GFW only in the noise condition among the ten-year-olds. The t-values and their significance are presented in Table 3. The apparent disparity between the magnitudes of the differences between some pairs of mean scores and the significance level of the t-values obtained is a product of the differences in the standard deviations of performances across the various categories. The significant t-values suggest the presence of a variable affecting performance associated with repeated presentations of speech discrimination tests under the specified age group and listening condition categories. These data indicate that repeated presentations can have a significant effect on scores for the five-year-olds under

TABLE 3. T-values, degrees of freedom and significance levels for differences between mean GFW and WIPI repeated and single presentation speech discrimination percentage scores of five-year-old group at 40 dB SL and ten-year-old group at 24 dB SL under quiet and noise listening conditions.

	Repeated Presentation		Single Presentation				
	Mean	SD	Mean	SD	t	p	df
Five-Year-Olds							
Quiet WIPI	98.0	5.1	91.2	6.2	2.68	.015	18
Quiet GFW	98.0	3.6	87.0	6.8	4.56	<.000	18
Noise WIPI	90.0	7.8	76.4	9.9	3.41	.003	18
Noise GFW	86.0	12.0	77.4	11.7	1.63	.121	18
Ten-Year-Olds							
Quiet WIPI	99.6	1.3	95.6	4.8	2.55	.020	18
Quiet GFW	97.3	2.6	94.0	4.1	2.16	.045	18
Noise WIPI	93.6	4.7	88.0	4.2	2.81	.012	18
Noise GFW	96.7	5.0	87.7	5.0	4.04	.001	18

both listening conditions. The effect of repeated presentations is apparent for the ten-year-olds only when background cafeteria noise is present as a distractor.

The mean scores for both tests under each listening condition were plotted to delineate the performance-by-intensity curves. Different curves were drawn to represent the performance of each of the two age groups in response to each of the discrimination tests. Figure 1 represents the performances of the five-year-olds. Figure 2 represents the performances of the ten-year-olds. Each figure has two parts, one representing each listening condition. Figure 3 presents the PI curves derived from the performances in the absence and in the presence of background cafeteria noise. The coordinate systems of Figure 3 contain the performance curves of both age groups so that inter-age group comparisons may be visualized conveniently. Figure 3 is arranged in two parts, one for each listening condition.

In general, the curves demonstrate that performances on the GFW, even more so, and the WIPI speech discrimination tests improve at a relatively rapidly accelerated rate as the sensation level increases from 8 dB SL to 16 dB SL. Thereafter, the improvement in performances decelerates and tends toward a plateau. There does not appear to be any evidence which indicates particular differences between the two tests based on their PI curves.

The slopes of the PI curves between the 8 and 16 dB sensation levels represent the acceleration of the line

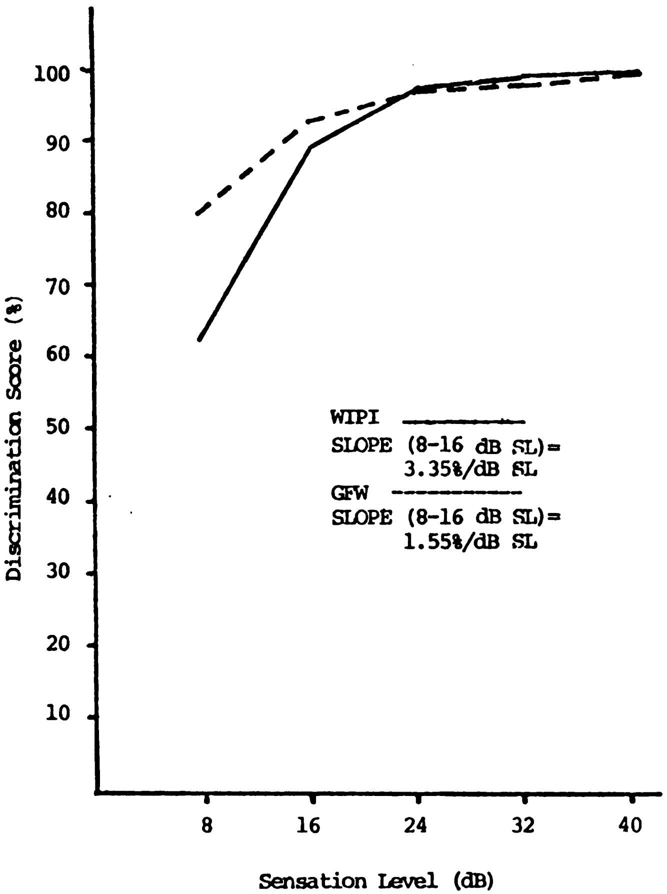

FIGURE 1a. Performance by intensity functions for five-year-olds' performances on the WIPI and GFW speech discrimination tests in the quiet listening condition.

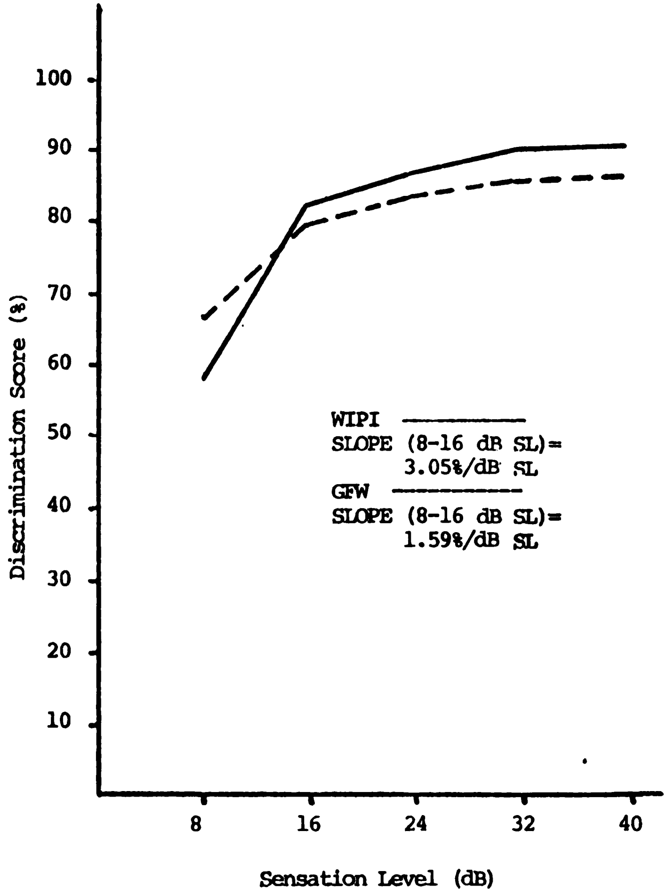

FIGURE 1b. Performance by intensity functions for five-year-olds' performances on the WIPI and GFW speech discrimination tests in the presence of background cafeteria noise.

FIGURE 1

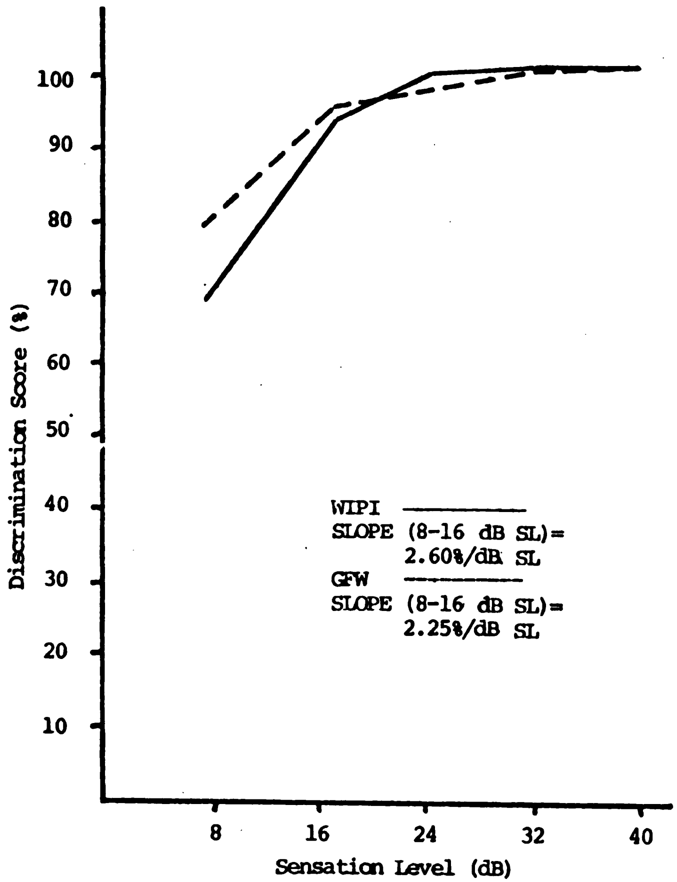

FIGURE 2a. Performance by intensity functions for ten-year-olds' performances on the WIPI and GFW speech discrimination tests in the quiet listening condition.

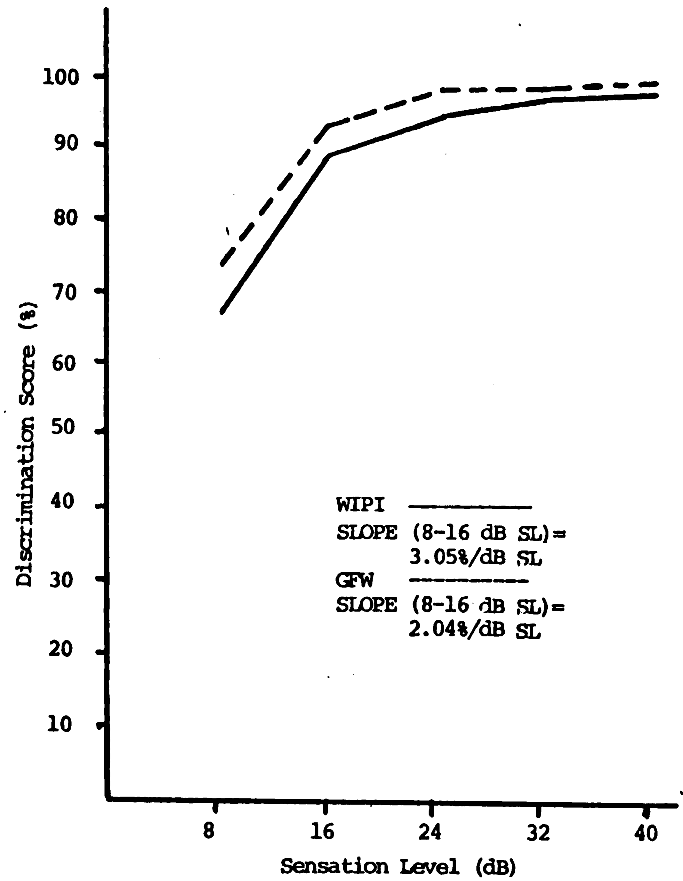

FIGURE 2b. Performance by intensity functions for ten-year-olds' performances on the WIPI and GFW speech discrimination tests in the presence of background cafeteria noise.

FIGURE 2

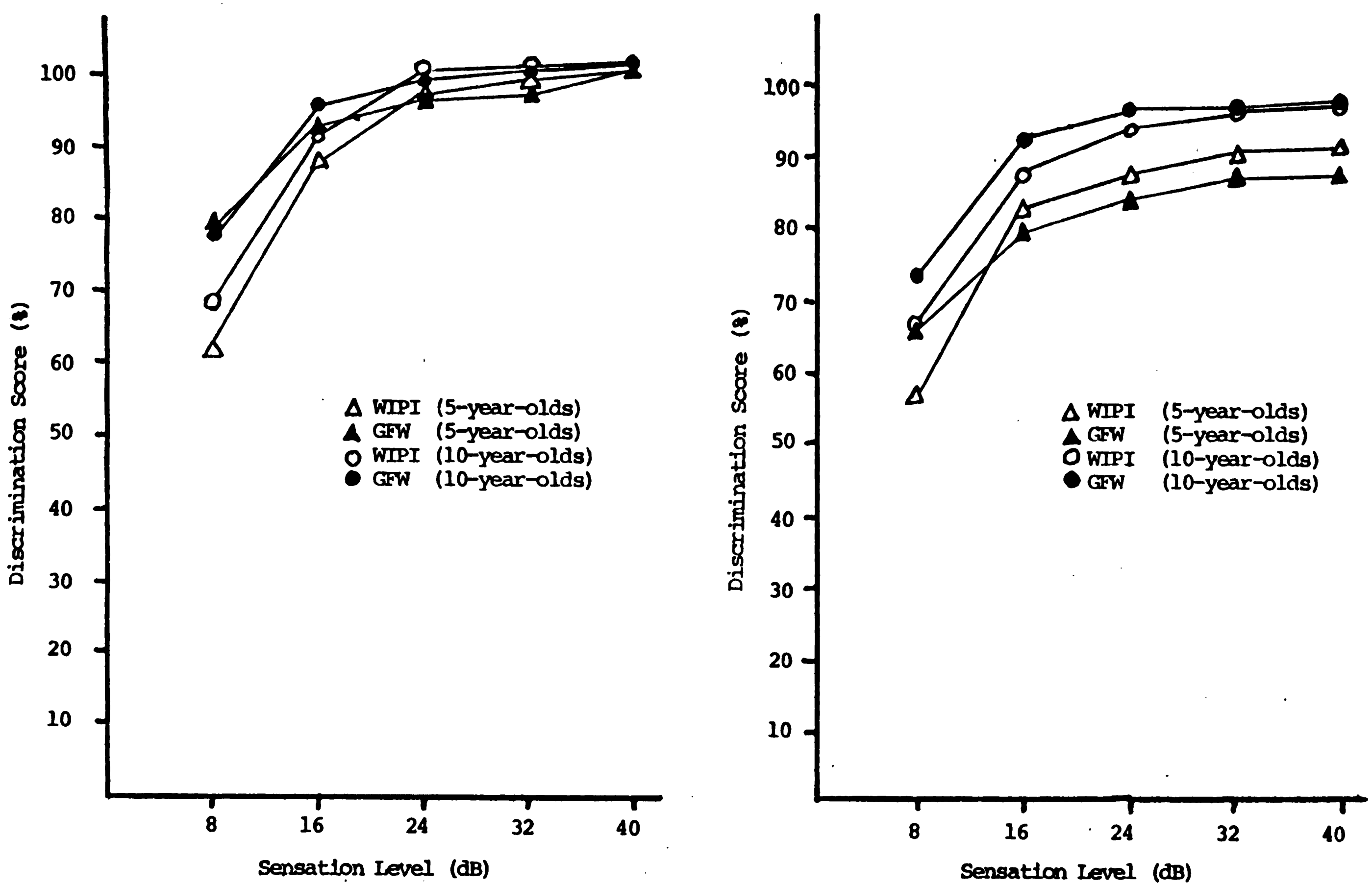

Figure 3a. Performance-by-intensity functions for five and ten-year-olds' performances on the WIPI and GFW speech discrimination lists in the quiet listening condition.

Figure 3b. Performance-by-intensity functions for five and ten-year-olds' performances on the WIPI and GFW speech discrimination lists in the presence of background cafeteria noise.

FIGURE 3

connecting the two points (Figures 1 and 2). The computational formula used to determine the slopes (Person, 1973) was as follows:

$$\text{Slope (\%/dB SL)} = \frac{\text{Score A\% (@16 dB SL) - Score B\% (@8 dB SL)}}{\text{16 dB SL - 8 dB SL}}$$

Differences in the magnitudes of the slope values fell short of the two percent/dB criterion in all cases. Overall, the slopes, measured in percent per decibel, for the GFW were lower in magnitude than those for the WIPI, suggesting a more gradual increase in performance for the GFW as sensation level is increased.

A repeated measures analysis of variance (Winter, 1971) permitted examination of the performances of all the repeated presentation subjects. Table 4 presents a summary of this analysis.

Those F-values which were significant ($p = .01$) between subjects were those associated with the five-year-olds' versus ten-year-olds' performances and quiet versus noise listening conditions. These data suggest that, overall, the ten-year-olds' performances on the two tests of speech discrimination in both listening conditions are superior to those of the five-year-olds. Further, the significance of the quiet versus noise F-value suggests

TABLE 4. Summary of repeated measures analysis of variance for all subjects in repeated presentation speech discrimination test groups.

SOURCE OF VARIATION	S.S.	d.f.	M.S.	F
BETWEEN SUBJECTS	23,130.52	39		
FIVE- vs. TEN-YEAR-OLDS	3,803.80	1	3,803.80	9.98*
QUIET vs. NOISE	4,837.90	1	4,837.90	12.69*
AGE vs. LISTENING CONDITION	761.48	1	761.48	2.00
ERROR BETWEEN GROUPS	13,727.34	36	381.32	
WITHIN SUBJECTS	50,640.43	360		
SENSATION LEVEL (WIPI and GFW)	40,161.32	9	4,462.32	4.23*
AGE vs. SENSATION LEVEL	444.30	9	49.37	0.05
LISTENING CONDITION vs. SENSATION LEVEL	117.20	9	13.02	0.01
AGE vs. LISTENING CONDITION vs. SENSATION LEVEL	429.75	9	47.75	0.05
ERROR WITHIN GROUPS	9,487.86	328	1,054.21	

* $F_{.99}$ (1, 36) between subjects = 7.39

$F_{.99}$ (9, 328) within subjects = 2.46

that performances by all the children are better under quiet conditions than they are in noise conditions.

The F-value which was significant ($p \leq .01$) within subjects was related to sensation level changes. This significant F-value indicates that at least one of the intensity increases caused a significant increase in performances by both age groups under both listening conditions. This implication is consistent with the general trend apparent by examination of the PI curves discussed earlier. The repeated measures analysis of variance alone does not indicate which of the sensation level increases brought about significant increases in performances. A Neuman-Keuls *a posteriori* treatment identified the sensation levels which elicited significant improvements in the subjects' scores. The magnitudes of these sensation levels varied between age groups and listening conditions. A discussion of the relative significance of each sensation level increase appears in the following section.

Listening Condition: Quiet Versus Noise Performance

A separate repeated measures analysis of variance permitted examination of the question regarding the effect of noise on the children's performances, using the scores obtained with the repeated presentation groups under each listening condition. This analysis was also appropriate since the F-value related to listening condition obtained with the overall analysis was significant ($p \leq .01$). Tables 5 and 6 present a summary of each repeated measures analysis of variance.

For the quiet listening condition, F-values were significant only within subjects. The F-value related to changes in sensation level was significant, suggesting, again as expected, that changes in presentation level accompanied changes in performance. The F-value for test versus sensation level indicates that the relationship between performance and sensation level is consistent for both the GFW and the WIPI.

The improvement in mean discrimination scores accompanying the increases in presentation level shows a rapid acceleration at the lower sensation levels, followed by a gradual deceleration as the mean scores approach or reach the 100 percent upper limit. The PI curves for both age groups in the quiet listening condition indicate that both tests are highly similar in performance-by-intensity functions at the higher sensation levels. Along these lines, it is of particular interest to note the lack of significance of the F-values related to the difference between the GFW and the WIPI as well as between the two age groups. These data suggest that, under quiet listening conditions, there was no significant difference in the performances of this experimental group related to which of the two tests was taken. Further, there appears to be no difference between

TABLE 5. Summary of repeated measures analysis of variance for all subjects in repeated presentation speech discrimination test groups under quiet listening conditions.

SOURCE OF VARIATION	S.S	d.f.	M.S.	F
BETWEEN SUBJECTS	7,450.19	39		
WIPI vs. GFW	417.02	1	417.02	2.28
FIVE- vs. TEN-YEAR-OLDS	405.55	1	405.55	2.22
TEST vs. AGE GROUP	41.60	1	41.60	0.23
ERROR BETWEEN GROUPS	6,586.02	36	182.95	
WITHIN SUBJECTS	29,641.85	160		
SENSATION LEVEL	20,680.80	4	5,170.20	103.67*
TEST vs. SENSATION LEVEL	1,526.99	4	381.75	7.65*
AGE GROUP vs. SENSATION LEVEL	21.88	4	5.47	0.11
TEST vs. AGE GROUP vs. SENSATION LEVEL	230.50	4	57.63	1.16
ERROR WITHIN GROUPS	7,181.68	144	49.87	

* $F_{.99}$ (1, 36) between subjects = 7.39

$F_{.99}$ (4, 144) within subjects = 3.44

TABLE 6. Summary of repeated measures analysis of variance for all subjects in repeated presentation test groups in the presence of background cafeteria noise at S/N = 9 dB.

SOURCE OF VARIATION	S.S.	d.f.	M.S.	F
BETWEEN SUBJECTS	15,728.18	39		
WIPI vs. GFW	51.71	1	51.71	0.16
FIVE- vs. TEN-YEAR-OLDS	3,984.56	1	3,984.56	12.53*
TEST vs. AGE GROUP	243.55	1	243.55	0.77
ERROR BETWEEN GROUPS	11,448.36	36	318.01	
WITHIN SUBJECTS	25,375.62	160		
SENSATION LEVEL	19,312.02	4	4,828.01	130.63*
TEST vs. SENSATION LEVEL	575.19	4	143.80	3.89*
AGE GROUP vs. SENSATION LEVEL	26.66	4	6.67	0.18
TEST vs. AGE GROUP vs. SENSATION LEVEL	139.57	4	34.89	0.94
ERROR WITHIN GROUPS	5,322.18	144	39.96	

* $F_{.99}$ (1, 36) between subjects = 7.39

$F_{.99}$ (4, 144) within subjects = 3.44

the performances of these children in the quiet listening condition, regardless of which age group they were in.

The emergence of a significant F-value related to sensation level increases led to the investigation of which specific sensation level changes produced the significant effects. A Neuman-Keuls *a posteriori* procedure permitted a test for the significance of the differences between scores for the various sensation levels. Table 7 presents a summary of the Neuman-Keuls treatment for all the experimental groups' speech discrimination scores in the quiet listening condition. The Neuman-Keuls treatment results indicate that the discrimination scores in quiet are significantly different ($p \leq .01$) at each sensation level until 24 dB SL is reached. Thereafter, the differences cease to be significant. There is no evidence of significance between the 32 dB sensation level and the 24 dB sensation level. The implication from these data is that although the magnitude of the speech discrimination scores in the quiet listening condition continues to increase, the amount of increase at each sensation level beyond 24 dB is not significant.

The presence of cafeteria noise as a background distraction appeared to bring about changes in the subjects' performances. In other words, the initial impression gained from an examination of the means and standard deviations of all the subjects is that the WIPI and GFW scores are poorer in the presence of background noise than they are in quiet.

TABLE 7. Summary of Neuman-Keuls <u>a posteriori</u> treatment for differences in speech discrimination scores of all subjects receiving repeated test presentations in quiet listening condition. (* indicates significance at or below .01 level.)

TRUNCATED RANGE r	2	3	4	5
$q_{.99}(r,\infty)$	3.64	4.12	4.40	4.60
$q_{.99}(r,\infty)\sqrt{nMSerror}$	162.71	184.01	196.65	205.58

SENSATION LEVELS (dB)	8	16	24	32	40
8		*	*	*	*
16			*	*	*
24					
32					
40					

The overall repeated measures analysis of variance substantiates this impression with a significant F-value ($p \leq .01$) between subjects for the listening condition variable (see Table 4, p. 90).

A separate repeated measures analysis of variance treatment for the speech discrimination scores in the noise listening condition resulted in the emergence of a significant F-value ($p \leq .01$) for age group differences between subjects. Within subjects, the F-value related to sensation level increases was significant ($p \leq .01$), as was an additional F-value relating sensation level changes to increases in both the WIPI and GFW scores. Table 6 gives a summary of the noise condition repeated measures analysis of variance.

Inspection of the PI curves (see Figures 1 and 2) for the noise listening condition reveals that, similar to the quiet condition, the magnitude of the discrimination scores increases with increases in sensation level. Since the scores appeared to plateau at the higher sensation levels, it was necessary to investigate the relative significance of differences between scores at each sensation level, in a manner similar to that used for the quiet listening condition data.

Table 8 presents a summary of the Neuman-Keuls treatment for determining the significance of differences ($p \leq .01$) between the scores at each sensation level. This treatment differed in one respect from the treatment used

TABLE 8. Summary of Neuman-Keuls a posteriori treatment for differences in speech discrimination scores of all subjects receiving repeated test presentations in the presence of cafeteria background noise at S/N = 9 dB. (* indicates significance at or below the .01 level.)

TRUNCATED RANGE r	2	3	4	5	6	7	8	9	10
$q_{.99}(r,\infty)$	3.64	4.12	4.40	4.60	4.76	4.88	4.99	5.08	5.16
$q_{.99}(r,\infty)\sqrt{nMSerror}$	98.97	112.02	119.63	125.07	129.42	132.68	135.67	138.12	140.29

SENSATION LEVELS (AGE GROUP)	8(5)	8(10)	16(5)	24(5)	32(5)	40(5)	16(10)	24(10)	32(10)	40(10)
8(5)		*	*	*	*	*	*	*	*	*
8(10)			*	*	*	*	*	*	*	*
16(5)						*	*	*	*	*
24(5)								*	*	*
32(5)								*	*	*
40(5)								*	*	*
16(10)									*	*
24(10)										
32(10)										
40(10)										

in the quiet listening condition. Since the F-value related to age was significant, it was appropriate to examine the differences between age groups at each sensation level. A discussion of differences noted which related specifically to age is under a separate heading.

Examination of the differences between discrimination scores obtained at the five sensation levels in the presence of background cafeteria noise reveals a trend different from that observed for the quiet condition performances. Each corresponding increase in the mean performance scores failed to represent a significant improvement ($p \leq .01$) over the previous level's score as the sensation level increased beyond 6 dB. This trend continued until the sensation level of the stimuli reached 40 dB. A significant increase in scores emerged at this sensation level for the five-year-old group only. For the ten-year-old group, a significant increase in performance over the previous level's performance did not accompany the final 8 dB incremental increase in stimulus presentation level.

Age Effects: Five Versus Ten-Year-Old Performance

The overall repeated measures analysis of variance indicated that age was a significant variable ($p \leq .01$) among the experimental subjects (see Table 4). Tables 1 and 2, which list the mean discrimination scores for the two age groups, show that, in general, the ten-year-old group

achieved higher scores than did the five-year-old group. The findings of the repeated measures analysis of variance and Neuman-Keuls treatment performed on the data obtained with the group listening under the noise condition as a separate entity substantiates this impression. The repeated measures analysis of variance of the data obtained under the quiet listening condition only failed to yield a significant F-value ($p \leq .01$) for age. This lack of significance is of some interest because it suggests that, for this experimental group, there was no significant difference between the performances of five-year-olds and those of ten-year-olds on the WIPI and the GFW in the absence of a noise distraction.

Repeated measures analysis of the data obtained under the noise condition only resulted in the between subjects variable of age group emerging as significant ($p \leq .01$). Table 6 presents a summary of the repeated measures analysis of variance. Within subjects, the findings were similar to those obtained with the data collected under quiet conditions. Sensation level changes yielded a significant ($p \leq .01$) F-value. Another significant F-value ($p \leq .01$) indicated that the sensation level changes were consistent for both the WIPI and the GFW.

The emergence of the significant F-value for age suggested that in noise, sensation level changes may not have had equal effects on both age groups. For this reason, it was appropriate to examine the differences between the

speech discrimination scores in a manner that took into account age group as well as sensation level changes. A slight modification of the Neuman-Keuls *a posteriori* treatment accomplished this task.

Table 8 summarizes the results of the Neuman-Keuls treatment for the data obtained under the noise listening condition. As mentioned in the previous section, the treatment used for the data obtained under the noise listening condition compared the scores of each age group at each of the different sensation levels.

Inspection of the Neuman-Keuls treatment of the scores obtained in the noise condition reveals some differences between the performances of the two age groups. All these differences were significant at the .01 level. Significant differences included variations in the way intensity increases affected scores of the two age groups. Significant differences also emerged in the overall magnitude of the scores of the five-year-olds as compared to those of the ten-year-olds.

The effect of increases in stimulus sensation level with the cafeteria noise distraction was different for each age group. Both age groups showed significant increases ($p \leq .01$) in scores accompanying the first 8 dB increase in sensation level. Each subsequent increase in sensation level thereafter failed to produce a significant increase in performance over that of the previous level. Among the

five-year-old group, this trend continued until the stimulus intensity reached a critical sensation level of 40 dB, whereupon a significant increase in performance appeared.

For the ten-year-olds, scores at both the 32 and 40 dB sensation levels were significantly better than scores elicited at the 16 dB sensation level. For the five-year-olds, there was no significant difference in scores between the 32 and the 16 dB sensation levels, whereas, the difference between these levels was significant for the ten-year-old group. However, increasing the sensation level from 32 to 40 dB did not bring about a significant performance improvement among the ten-year-olds. These data imply that the increase in stimulus intensity from 16 to 32 dB SL in two 8 dB steps brought about a significant increase in scores for ten-year-olds but not for five-year-olds. The five-year-olds apparently needed an increase in signal intensity from 16 to 40 dB SL in three 8 dB steps in order to increase their speech discrimination scores to a similar extent.

Significant differences were also noted between the magnitudes of the five-year-olds' scores and the magnitudes of the scores of the ten-year-olds. Table 8 shows that the ten-year-old group scores are significantly higher ($p \leq .01$) than those of the five-year-old group at each of the five sensation levels in the presence of background cafeteria noise. The scores of the ten-year-old group under this

listening condition at the 24 dB sensation level were also significantly greater than those made by the five-year-old group at both the 32 dB and 40 dB sensation levels. These findings have two implications. First, the ten-year-olds were better performers on the WIPI and the GFW at each sensation level with a noise distraction than the five-year-olds. Second, when a noise distraction was present, the ten-year-olds were better performers at 24 dB SL than the five-year-olds were at 40 dB SL. This second observation was consistent for both the WIPI and the GFW speech discrimination tests.

T-values were calculated for the single presentation groups to determine if the effects of age were significant at the .01 level. In the quiet listening condition, the difference between the five- and ten-year-old groups was significant for the GFW only ($t = 2.79$; $df = 18$). In the noise listening condition, significant differences emerged between the two age groups for both the GFW ($t = 2.55$; $df = 18$) and the WIPI ($t = 3.39$; $df = 18$). The data for the single presentation groups appears to contrast with those for the repeated presentation groups in that age related performance differences exist for the GFW in the quiet listening condition.

WIPI and GFW Correlations

Correlation coefficients were computed for comparison of the WIPI and GFW and are presented in Table 9.

TABLE 9. Pearson product-moment correlation coefficients for comparison of WIPI and GFW speech discrimination scores of all subjects receiving repeated presentations in quiet and noise listening conditions. (* indicates significance at or below the .01 level. A dash indicates that correlation coefficient cannot be computed.)

FIVE-YEAR-OLDS		
SENSATION LEVEL (dB)	CORRELATION COEFFICIENT (QUIET)	CORRELATION COEFFICIENT (NOISE)
8	.615	.591
16	.657	.744*
24	.804*	.542
32	.876*	.458
40	.322	.739

TEN-YEAR-OLDS		
SENSATION LEVEL (dB)	CORRELATION COEFFICIENT (QUIET)	CORRELATION COEFFICIENT (NOISE)
8	.376	.593
16	.362	.119
24	.085	.508
32	-	.610
40	-	.511

For the five-year-old age group, the correlation coefficients reached significance at the $p \leq .01$ level only for the 24 dB and 32 dB sensation levels in the quiet listening condition. Significance was reached only at the 16 dB sensation level in noise. For the ten-year-old group the correlation coefficients never achieved significance at the .01 level of confidence.

At the two highest sensation levels in quiet, the coefficients could not be computed for mathematical reasons. The significance of the correlation coefficients belies the relationship between the scores in analyzing speech discrimination test results because of the dependence of such coefficients on the variation of the scores (Thornton and Raffin, 1978). If the variation of the group of scores is zero, as was the case for the ten-year-old group in the quiet listening condition at the 40 dB sensation level, the coefficient cannot be computed. Correlation coefficients are measures of predictability of changes in scores. As such, they tend to lose meaning as comparative measures when the variation is very small or nonexistent.

Summary

This chapter presented the performance-by-intensity functions of five- and ten-year-old children on two multiple choice picture tests of speech discrimination under quiet and noisy listening conditions. These data suggest that changes in performance on the WIPI and GFW tests of speech

discrimination accompany changes in three other variables: stimulus sensation level; quiet versus noisy listening conditions; and age group of the subjects.

Increases in performances of the experimental subjects resulted from four 8 dB increases in the sensation level of the stimuli. In the quiet listening condition, this increase in performances was large at first but diminished as the sensation level increased at each subsequent step. In the presence of background cafeteria noise, a large performance increase occurred at first, but over subsequent trials more intensity was needed to significantly increase the speech discrimination scores than was needed for the quiet listening condition.

It was apparent that the repeated presentation of the GFW and/or the WIPI resulted in higher scores under some conditions than would have been obtained with a single presentation of the test under the same condition.

Differences associated with age were also apparent but only in the presence of background cafeteria noise. The ten-year-olds under study seemed to perform significantly better than the five-year-olds at all the sensation levels when distracting background noise was present. No differences in performance emerged between the two age groups in the quiet listening condition.

The presentation and analysis of the data contained in this chapter represent answers to the five research

questions of this study. There was no significant difference ($p \leq .01$) between the mean performance-by-intensity functions obtained with the WIPI and the GFW for either the five-year-old or the ten-year-old age group. The presence of background cafeteria noise had a significant effect ($p \leq .01$) on the performances of children in both age groups. Age group was a significant variable ($p \leq .01$) related to listener performance. Pearson product-moment correlation coefficients were unsatisfactory as a means of inter-test score prediction. Significant differences ($p \leq .01$) were found between the performances of children who received a single presentation of the speech discrimination tests and the performances of the children who received multiple presentations. This finding was consistent for both age groups and both listening conditions.

Chapter 5

SUMMARY AND CONCLUSIONS

Summary

This study had four stated purposes: first, to compare the performances of two groups of normal hearing children as a function of intensity in a sound field on two multiple choice-type picture tests of speech discrimination, the WIPI and the GFW; second, to identify an optimum intensity in decibels above the speech reception threshold at which these tests should be presented in the sound field to children represented by the sample studied herein; third, to identify the possible existence of a significant difference in performance on these tests between two age groups, five and ten years, in the presence of background cafeteria noise; and, fourth, to examine the effects of multiple test presentations upon listener performance.

The importance of this study was based on three premises. First, since little information was available concerning performance-by-intensity measurement with children on multiple-choice-type picture tests of speech discrimination, it appeared of importance to both the speech and language pathologist as well as the audiologist to present results pertaining to how normal hearing children in two age groups perform on these tests when stimulus intensity and listening condition are manipulated. Second, it might well be relevant to the evaluation and treatment of certain speech and language disorders to maintain control of stimulus intensity when speech discrimination tests are administered in a sound field situation. A third basis of this study's importance lay in the fact that the WIPI and the GFW performances may be commonly compared by audiologists and speech and language pathologists who are interested in referrals of clients to one another. Finally, it appeared important to include the GFW as a speech discrimination list to be compared to other speech discrimination lists, in particular, the WIPI, so that audiologists might be provided with a sound basis for choosing the GFW as an alternative multiple-choice type picture test of speech discrimination.

The following research hypotheses were drawn for the purposes of this study. They are stated here in null hypothesis form.

1. The differences between the slopes of the portions of the mean PI function curves between 8 and 16 dB SL for the two tests within each age group, in both quiet and noise conditions, will not be more than 2%/dB SL.
2. The point of 90% or better correct discrimination for each subject will not occur at the same sensation level for both tests within each age group in the quiet listening condition.

3. The difference between the slopes of the portions of the mean PI function curves between 8 and 16 dB SL for both tests between age groups will not be greater than 2%/dB SL.
4. There will be no difference between the minimum sensation level at which 90% or better correct discrimination occurs for each subject in the quiet listening condition between age groups for both tests.
5. Correlation coefficients between scores on the two tests within each age group will not be significant at the $p \leq .01$ level for both listening conditions.
6. Differences between discrimination scores related to age will not be significant at the $p \leq .01$ level at each sensation level.
7. Differences between discrimination scores related to quiet versus noise listening conditions will not be significant at the $p \leq .01$ level.
8. There will be no difference ($p \leq .01$) between the mean speech discrimination scores of the experimental group that hears repeated presentations of the two tests and those of the group which hears only one presentation of the test at a specified sensation level.

The review of the literature related to this study made an attempt to point out that semantic differences exist between the terms "auditory discrimination," "speech discrimination" and "word discrimination."

The literature reviewed portrayed speech discrimination, referring to the differential perception of speech sounds, as a developmental phenomenon. This development begins in earliest infancy and rapidly improves to eleven or twelve years of age. After this age, further improvement accompanying age appears to become less and less remarkable. In old age, various degrees of decrement in speech discrimination ability were documented.

A relationship has been documented between speech discrimination abilities and other variables such as speech articulation, intelligence and minimal brain dysfunction. Specifically, poor performance on speech discrimination tests accompanied the more sizable speech articulation disorders. Children of low intelligence performed poorer on speech discrimination tests. Children with diagnosed minimal brain dysfunction performed poorer than children free of such diagnoses on tasks which required auditory figure-ground perception. Dialect and environment were also identified as important factors in speech discrimination test performance. Children appear to be more efficient speech discriminators if the speech signal matches their own dialect. Noisy, unhealthy environments were associated with diminished speech discrimination abilities.

Speech discrimination test development was traced from its inception in the early 1930's to the present time, with variations in stimulus presentation and listener requirements being identified as factors affecting performance. Changes in listener performance accompanied changes in the speaker, the nature of the stimulus (whether word or unmeaningful syllable), and changes in task requirements (open versus closed set; written versus spoken response). The listener's vocabulary also affected speech discrimination performance. The WIPI and the GFW, two multiple-choice-type picture tests of speech discrimination, were described in detail as the instruments used in the present investigation.

Finally, the second chapter discussed the literature related to the performance-by-intensity function as a means of comparing speech discrimination word lists and as a useful clinical differential diagnostic method. The literature demonstrated that the performance-by-intensity function (abbreviated PI function) is a useful tool for the differential diagnosis of auditory dysfunction. Several studies made use of the PI function to compare speech discrimination test instruments.

There were eighty subjects in this study. Forty of these were age five and forty were age ten. Within each age group, subjects were randomly assigned to either the quiet or the noise listening condition. Half of the subjects were administered the GFW and the WIPI, in a counterbalanced

design, at five equal ascending sensation levels from 8 dB SL to 40 dB SL. Each subject in the remaining half received a single presentation of both tests, in a counter-balanced design, as a means of examining the effects of repeated presentation of the word lists. This single presentation group received both word lists at the sensation level at which each of the subjects exposed to the repeated presentations achieved a score of 90% or better or at the highest (40 dB) sensation level tested if all the subjects exposed to repeated presentations did not reach this criterion.

All the subjects were screened prior to the speech discrimination testing to assure a normal peripheral auditory mechanism. Each subject passed a pure tone hearing screening at 20 dB HTL, was free of a history of chronic middle ear pathology and frequent exposure to high intensity noise. All of the subjects spoke General American English Dialect.

The speech discrimination stimuli were administered according to the test manual instructions, except that a special stimulus tape recording was used for this project. All speech discrimination stimuli were delivered in a sound field via an Intensify model 1001-A loudspeaker. Intensity was controlled with a Grason-Stadler model 1701 clinical audiometer. Subjects were tested in a sound treated booth. For the noise listening condition, the signal-to-noise ratio was held at a constant 9 dB. Cafeteria noise was

used as a distraction in the noise listening condition as in the standard GFW format.

Calculation of the mean speech discrimination scores for all the subjects was the first step in the data analysis. For the subjects listening at five sensation levels, a graphic representation of these scores depicted the relationship of test performance to sensation level for each age group and listening condition. Three separate repeated measures analysis of variance treatments performed on the speech discrimination scores as a whole and upon those obtained under each listening condition separately allowed statistical interpretation of the relationships of test performance to age group, listening condition and sensation level. Pearson product-moment correlation coefficients computed for the two tests at each sensation level represented an attempt to measure inter-test score predictability. Two-tailed t-tests applied to test the significance of the differences between the speech discrimination scores of the subjects who heard repeated presentations of the two word lists and those who heard only one presentaion of the two word lists allowed statistical interpretation of these differences.

Conclusions

The following conclusions are appropriate within the limitations of this study. Their order of presentation corresponds to that of the hypotheses presented earlier in this chapter.

1. The GFW and the WIPI are equivalent measures of speech discrimination for children from populations represented by the samples studied herein. The differences between the slopes of the portions of the PI curves of the GFW and the WIPI were less than 2%/dB SL whether the listening condition involved was noisy or quiet. These results indicate a failure to reject the first null hypothesis. The small amount of difference for slope is one factor which suggests that the GFW and the WIPI are equivalent measures of speech discrimination when presented under the same conditions.

The lack of significant ($p \leq .01$) differences between the GFW and WIPI scores of the children within each age group and listening condition emerges as further data to support the conclusion of equivalence between the two tests as long as the presentation specifications remain constant.

The fact that the test training procedures, recommended by the authors of the test manuals, differ somewhat in concept mitigates interpretation of these data. The GFW authors designed their training procedure to assure the examiner that his or her client knows the test requirements as well as the vocabulary involved in the word list. The WIPI test training procedure, on the other hand, simply assures that the client is aware of the picture pointing task requirements and will comply.

The conclusions resulting from the present study suggest that training of only the GFW vocabulary is a

prerequisite for obtaining equivalent results between the GFW and the WIPI, since the experimenter followed each test's specified training procedure. A further conclusion is that, in the event that a client is untrainable to the GFW vocabulary requirements, results which are substantially the same may be obtained by using the WIPI, as long as the stimulus presentation specifications are not changed.

Differences exist between the WIPI and GFW formats. A subject responding to GFW stimuli had four pictures from which to choose. Thus, there was a 25% chance of producing a correct response through guessing. This chance was reduced to 17% for the WIPI, since the subjects' task was to choose the correct picture among five foils. The phonetic nature of the words represented by the alternative pictures on each plate also differed between the two tests. Whereas, for the GFW, the four test pictures represented minimal pair contrasts (i.e., words differing by only one phoneme), the WIPI showed no consistent phonetic arrangement. Out of the twenty-five WIPI test plates, only three represented minimal pair contrasts. Four of the WIPI test plates presented vowel contrasts as well as consonant contrasts on the same test plate. These factors may also have contributed to the similarity of the subjects' performances on the two tests in spite of the training procedure differences.

The holding constant of stimulus presentation characteristics precludes the use of prerecorded tape recordings provided by the different test publishers to obtain

equivalence between GFW and WIPI test results for the present. The effects of differences between such differently produced tape recordings on children's performances are yet unknown. The results of Kruel, Bell and Nixon's (1969) research suggest the possibility of such differences.

2. There was no difference in the sensation levels at which 90% or better correct discrimination occurred for either test in the quiet listening condition within each age group among the repeated presentation groups. Thus, the second null hypothesis is rejected. This fact indirectly supports the conclusion of equivalence between the GFW and the WIPI when presented under the same conditions with the same presentation technique. More importantly, this fact suggests clinically useful minimum sensation levels when examining PI-PB functions for possible auditory disorders among children of appropriate ages.

The authors of the GFW presented percentile scores obtained with various age groups in their standardization study (Goldman, Fristoe and Woodcock, 1971). They suggested a "cutting score" at the 20th to 30th percentile as a useful index when employing a subject's performance on the GFW as a basis for referral to an audiologist. The mean scores of all groups of subjects in the present study exceeded this criterion in all cases. Percentile scores for the single presentation groups in the quiet listening condition were somewhat lower than those of their repeated presentation counterparts, and one five-year-old subject's

(B.S.) performance ranked at the 14th percentile for her age group. The absence of communication problems and normal receptive vocabulary in this child suggest that the GFW norms and "cutting score" should be interpreted in conjunction with other independent test data when evaluating a given child.

3. The differences between the slopes of the portions of the mean PI function curves between 8 and 16 dB SL for both tests between age groups were less than 2%/dB SL. These differences did not exceed 2%/dB SL in either listening condition. This indicates a failure to reject the third null hypothesis.

The small slope differences, together with the lack of significance of the age variable in the quiet condition data analysis suggest that the function of intensity as it affects performance on these two tests is substantially the same for both the five- and ten-year-old groups studied herein. Under noisy conditions, the data anlysis revealed that significant increases in the five-year-old group's scores resulted only after a greater sensation level increase that was necessary for the ten-year-old group. At low sensation levels, the effect of sensation level increases between age groups was quite similar. Figure 3 compares the PI functions for both age groups in each listening condition.

4. The data indicate rejection of the fourth null hypothesis. There was a difference between the minimum

sensation levels of 90% or better correct discrimination for each subject in the quiet listening condition between age groups. This sensation level was 24 dB for the ten-year-old group. For the five-year-old group, the 90% or better performance occurred at 40 dB SL with one exception. One of the five-year-old subjects failed to reach the 90% criterion in response to the WIPI at any sensation level. These data lead to the conclusion that ten-year-olds need less stimulus intensity in order to discriminate between words and point to pictures of them at a 90% accuracy level than do five-year-olds.

A comparison of the data obtained with both age groups as well as in both the single and repeated presentation groups suggests that there may be different minimum stimulus intensity levels necessary to elicit maximum performance on the WIPI and GFW from different age groups of children. Of the two age groups of children studied herein, the ten-year-olds with no prior experience with the test performed at a 90% or better level under quiet listening conditions and a stimulus presentation level of 16 dB. In the presence of background cafeteria noise, more intensity or more exposure to the test instrument was necessary. The five-year-olds studied herein needed at least 24 dB SL of stimulus intensity as well as some experience with the test instrument beyond the recommended training procedure for 90% or better success in quiet listening conditions. Under noisy listening conditions, it may be that one cannot

expect five-year-olds to exceed the 90% performance level even after repeated trial runs with the same word list and a stimulus intensity of 40 dB SL.

Thus, recommended sensation levels for both tests depend upon the age of the subject to be tested. For five-year-old children, the sensation level of 40 dB is recommended in quiet, and at least the same amount of stimulus intensity in noise. For ten-year-olds, a minimum sensation level of sixteen dB is recommended in quiet for 90% or better performances on both tests. In noise, 24 dB SL is the minimum recommended presentation level for ten-year-olds. Improved performance on both tests by both age groups of children should result from one or more trial runs with the instrument. These results indicate an optimal presentation level for both age groups combined is at least 40 dB SL.

These data have implications for the classroom teacher. It appears that if one expects five-year-old children to discriminate the teacher's speech, at least for single word utterances, the classroom must be quiet. The teacher must also bear in mind that the intensity of spoken messages to the students must be at least 40 dB above each child's speech reception threshold. Ten-year-olds can be expected to perform better than five-year-olds, but the presence of background noise can impair their discrimination as well. Poor classroom performance by an elementary school student may well be related to a problem in the acoustic characteristics of the signal the teacher sends.

5. Within the limitations of this study, correlation coefficients are poor measures of the similarity of speech discrimination test results because of their dependence on the variation of scores, especially at the higher sensation levels. Failure to reject the fifth null hypothesis at the $p \leq .01$ level evolves from the inability to compute all the correlation coefficients.

6. The presence of background cafeteria noise is important if differences in speech discrimination test performances associated with age are to be investigated. Differences between discrimination scores related to age between five- and ten-year-olds were significant only in the presence of background cafeteria noise at a signal-to-noise ratio of 9 dB among the repeated presentation groups. There was little difference in repeated presentation multiple-choice type picture discrimination test performance between five- and ten-year-olds in quiet listening conditions. While the single presentation group data indicated age related differences in quiet GFW performances, these data compared groups listening at differing sensation levels. In the presence of background cafeteria noise, the ten-year-olds appeared to be better performers. These results indicate failure to reject the sixth null hypothesis at the .01 level for the quiet listening condition, and rejection of the sixth null hypothesis at the .01 level for the condition in which background noise was present.

7. Differences in the performances of children on the GFW and the WIPI occur between quiet and noisy listening conditions. Children in the age groups studied were better performers in quiet listening conditions than they were in noisy listening conditions. These data indicate rejection of the seventh null hypothesis at the .01 level.

8. Better speech discrimination scores can be expected from five-year-old and ten-year-old children who have had immediate previous experience with the test format and presentation characteristics than from those who have not. Therefore, the results of a single presentation of a picture speech discrimination test are a poor basis for making habilitative placement or treatment decisions.

There appear to be differences in performances on the GFW and the WIPI accompanying the repeated presentation of these tests as opposed to a single presentation. These data indicate rejection of the eighth null hypothesis at the .01 level. The scores obtained after repeated presentations are better than those obtained from a single presentation. These differences do not appear to be related to memorization of the word lists for two reasons. First, several subjects actually performed worse on a given test after a sensation level increase. Second, the mean scores obtained on the WIPI at 24 dB compare favorably with those obtained by Sanderson-Leepa and Rintelmann (1976) in quiet at the same sensation levels, except that the subjects of the present study scored lower at 8 dB SL, the first level

tested. Sanderson-Leepa and Rintelmann used four different word lists for their four sensation levels. The differences observed herein are probably related to the subject's becoming familiar with the task and with the acoustic characteristics of the speaker's voice.

The data reported in this study support the contention that a period of familiarization may be required if one expects to elicit the maximum speech discrimination performance from a given child. Goldman, Fristoe and Woodcock (1971) provided a familiarization process to be included as a standard part of the administration of their test. This process was oriented specifically to the test list vocabulary, however, instead of some of the other test materials. The present data suggest that a child might need to become familiar with speaker characteristics and presentation equipment in addition to the list vocabulary. The familiarization process provided with the WIPI is intended to acquaint the child with the nature of the task required. In the administration of the WIPI, no attempt is made to familiarize the child with list vocabulary or test materials.

One limitation to the application of the present findings regarding test familiarization is based upon the intended application of the test results in a clinical diagnostic situation. Normative data supplied with the GFW are based on a single presentation of the stimuli. If the purpose of speech discrimination testing is to determine a child's performance in relation to the standardized data,

then a single presentation is indicated. If the purpose of the speech discrimination testing is to examine the maximum performance of a given child, then several trial runs with the testing materials are recommended. Muma (1978) stressed the importance of repeated testing to obtain an accurate impression of a child's language ability. Ingram (1977) emphasized the need for several speech articulation testing experiences as a means of gaining a more valid picture of a child's speech articulation skills. The results of the present study suggest that there is a need for repeated testing of speech discrimination in a given individual as well.

Recommendations for Further Study

Several possible future research projects are recommended based on the present findings.

The differences in effects on performance of the presence of various types of background noise distractions are of interest. White noise, pink noise and narrow band noise distractions could be studied and related to quiet listening condition performance.

Performances of the GFW and WIPI should be examined among a geriatric population with and without various neurological pathologies known to affect speech and language performance.

It would be interesting to compare GFW performances of children listening to the publisher-prepared tape versus a different tape with a different speaker.

Children having various auditory pathologies should be subjects for a performance-by-intensity function study using the GFW or the WIPI. This might help provide further clinically useful data on the characteristics of auditory disorders in children. If auditory dysfunctions account for a substantial portion of learning disabilities (Rechner and Wilson, 1967; Perozzi and Kunze, 1971), children so diagnosed would make interesting subjects for further research along the lines of the present study.

The study of the PI functions of children with various other communication disorders might be illuminating. Children with functional articulation disorders, children with dysfluent speech, and children with cognitive or linguistic delays might show differences in the PI function when compared to the children studied herein.

Finally, the effects of right and left ear differences on PI functions would be a subject worthy of study. Using both normal and pathological populations from which to draw samples, it might well be possible to observe differences related to the presence or absence of lesions at various retrocochlear sites, including the auditory cortex of either cerebral hemisphere.

LIST OF REFERENCES

LIST OF REFERENCES

American National Standard Specifications for Audiometers, ANSI publication 53.6-1969 (reaffirmed 1973) New York: American National Standards Institute (1970).

Aungst, F., Frick, J., Auditory discrimination ability and consistency of articulation of /r/. Journal of Speech and Hearing Disorders, 29, 76-85 (1964).

Bender, L., Visual Motor Gestalt Test, New York: American Orthopsychiatric Association (1946).

Berg, W., Habituation and dishabituation of cardiac orienting responses in awake, four-month-old infants. Unpublished doctoral dissertation, University of Wisconsin, Madison (1971).

Berlin, C., Dill, A., The effects of feedback and positive reinforcement on the Wepman auditory discrimination scores of lower class Negro and White children. Journal of Speech and Hearing Research, 10, 384-389 (1967).

Broadbent, D., Perception and Communication, New York: Pergamon Press (1958).

Broadbent, D., Attention and perception of speech. Scientific American Reprint, 467 (1962).

Butterfield, E., Cairns, G., Discussion summary - infant reception research, in Schiefelbusch, R., and Lloyd, L., (eds.), Language Perspectives - Acquisition, Retardation, and Intervention, Baltimore: University Park Press (1974).

Carhart, R., Individual differences in hearing for speech. Annals of Otology, Rhinology and Laryngology, 55, 233-266 (1946).

Cohen, J., Diehl, C., Relation of speech-sound discrimination ability to articulation-type speech defects. Journal of Speech and Hearing Disorders, 28, 187-190 (1963).

Cohen, S., Glass, D., Singer, J., Apartment noise, auditory discrimination, and reading ability in children. Journal of Experimental Social Psychology, 9, 407-422 (1973).

Costello, M., Evaluation of auditory behavior using the Flowers-Costello test of central auditory abilities. In R. Keith (Ed.), Central Auditory Dysfunction. New York: Grune and Stratton (1977).

Culbertson, W., White, S., A comparison of performance-by-intensity functions for two picture identification auditory discrimination tests. Msha, 16, 6-12 (1980).

Davis, H., The articulation area and the social adequacy index for hearing. Laryngoscope, 58, 761-778 (1948).

Davis, H., Silverman, S., Hearing and Deafness, New York: Holt, Rinehart and Winston (1970).

Deutch, C., Auditory discrimination and learning: social factors. Merrill-Palmer Quarterly, 10, 277-296 (1964).

Dickson, S., Differences between children who spontaneously outgrow and children who retain functional articulation errors. Journal of Speech and Hearing Research, 5, 263-271 (1962).

Dirks, D., Effects of hearing impairment on the auditory system. Carterett, E., and Friedman, M., (eds.), Handbook of Perception, Volume IV, New York: Academic Press (1978).

Dirks, D., Kamm, C., Bower, D., The use of performance-intensity function for diagnosis. Journal of Speech and Hearing Disorders, 42, 408-415 (1977).

Doll, E., (ed.), The Oseretsky Tests for Motor Proficiency, Minneapolis: Educational Test Bureau, Education Publications (1946).

Dunn, L., Peabody Picture Vocabulary Test, Circle Pines, Minn.: American Guidance Service, Inc. (1965).

Egan, J., Articulation-testing methods. Laryngoscope, 58, 955-991 (1948).

Eilers, P., Minifie, F., Fricative discrimination in early infancy. Journal of Speech and Hearing Research, 18, 158-167 (1975).

Eilers, R., Wilson, W., Moore, J., Developmental changes in speech discrimination in infants. Journal of Speech and Hearing Research, 20, 766-780 (1977).

Eimas, P., Auditory and linguistic processing of cues for place of articulation by infants. Perceptual Psychophysiology, 16, 513-521 (1974a).

Eimas, P., Auditory and phonetic coding of the cues for speech: discrimination of the /r-l/ distinction by young infants. Perceptual Psychophysiology, 18, 341-347 (1975a).

Eimas, P., Siqueland, E., Juszyck, P., Vigorito, J., Speech perception in infants. Science, 171, 303-306 (1971).

Fairbanks, G., Test of phonemic differentiation: the rhyme test. Journal of the Acoustical Society of America, 30, 596-600 (1958).

Farquhar, M., Bankson, N., The Farquhar-Bankson In-depth Test of Auditory Discrimination, Unpublished test (1972).

Flanagan, J., Difference limen for formant amplitude. Journal of Speech and Hearing Disorders, 22, 205-212 (1957).

Flowers, A., Costello, M., Small, V., Manual for Flowers-Costello Test of Central Auditory Abilities. Dearborn, Mich.: Perceptual Learning Systems (1970).

Gaeth, J., A study of phonemic regression associated with hearing loss. Unpublished doctoral dissertation, Northwestern University, Evanston, Illinois (1948).

Goetzinger, C., Word discrimination testing, in Katz, J. (ed.), Handbook of Clinical Audiology, Baltimore: Williams and Wilkins (1972).

Goldman, R., Fristoe, M., Woodcock, R., A new dimension in the assessment of speech sound discrimination. Journal of Learning Disorders, 4, 364-368 (1971).

Goldman, R., Fristoe, M., Woodcock, R., Goldman-Fristoe-Woodcock Test of Auditory Discrimination, Circle Pines, Minn.: American Guidance Service, Inc. (1971).

Goldstein, K., Language and Language Disturbances, New York: Grune and Stratton (1948).

Gottesman, R., Auditory discrimination ability in Negro dialect speaking children. Journal of Learning Disabilities, 5, 94-101 (1972).

Gray, H., Michael, W., Sklar, M., Auditory perceptual thresholds in brain-injured children. Journal of Speech and Hearing Research, 8, 49-56 (1968).

Hall, M., Auditory factors in functional articulatory speech defects. Journal of Experimental Education, 7, 110-132 (1938).

Haskins, H., A phonetically balanced test of speech discrimination for children. Unpublished master's thesis, Northwestern University, Evanston, Illinois (1949).

Hirsch, I. J., Davis, H., Silverman, S. R., Reynolds, E. G., Eldert, E., Benson, R. W., Development of materials for speech audiometry. Journal of Speech and Hearing Disorders, 17, 321-337 (1952).

Holm, V., Kunze, L., Effects of chronic otitis media on speech and language development. Pediatrics, 43, 833-839 (1969).

House, A., Fairbanks, G., The influence of consonant environment upon the secondary acoustic characteristics of vowels. Journal of the Acoustical Society of America, 25, 105-113 (1953).

House, A., Williams, C., Hecker, M., Kryter, K., Articulation testing methods: consonantal differentiation with a closed-response set. Journal of the Acoustical Society of America, 37, 158-166 (1965).

Ingram, D., Phonological Disability in Children, New York: Elsevier (1980).

Jerger, J., Jerger, S., Diagnostic significance of P.B. word functions. Archives of Otolaryngology, 93, 573-580 (1971).

Jerger, J., Jerger, S., Estimating speech reception threshold from PI/PB function. Archives of Otolaryngology, 102 (8), 487-496 (1976).

Jones, K., Studebaker, G., Performance of severely hearing impaired children on a closed response auditory discrimination test. Journal of Speech and Hearing Research, 17, 531-541 (1974).

King, K., Infants' perception of acoustic cues for the voiced alveolar stop in two vowel environments. Unpublished doctoral dissertation, Michigan State University, East Lansing, Michigan (1976).

Kirk, S., McCarthy, J., Kirk, W., Illinois Test of Psycholinguistic Abilities, Urbana, Illinois: University of Illinois Press (1968).

Konkle, D., Beasley, D., Bess, F., Intelligibility of time-altered speech in relation to chronological aging. Journal of Speech and Hearing Research, 20, 108-115 (1977).

Kronvall, E., Diehl, C., The relationship of auditory discrimination and articulatory defects of children with no known organic impairment. Journal of Speech and Hearing Disorders, 19, 335-338 (1954).

Kruel, E., Bell, D., Nixon, J., Factors affecting speech discrimination test difficulty. Journal of Speech and Hearing Research, 12, 270-280 (1969).

Labov, W., Cohen, P., Robins, C., A preliminary study of the structure of English used by Negro and Puerto Rican speakers in New York City, New York: Columbia University (1965).

Lapko, L., Bankson, N., Relationship between auditory discrimination, articulation stimulability, and consistency of misarticulation. Perceptual and Motor Skills, 40, 171-177 (1975).

Lasky, R., Syrdal-Lasky, A., Klein, D., VOT discrimination by four-to-six month old infants from Spanish environments. Journal of Experimental Child Psychology, 20, 215-225 (1975).

Lehiste, I., Peterson, G., Linguistic considerations in the study of speech intelligibility. Journal of the Acoustical Society of America, 31, 280-286 (1959).

Lewis, F., Kelly, L., Oral stereognosis and auditory discrimination by adults. Perceptual and Motor Skills, 38, 12-18 (1974).

Lisker, L., Abramson, A., A cross-language study of voicing in initial stops: acoustical measurements. Word, 20, 384-422 (1964).

Locke, J., The inference of speech perception in the phonologically disordered child. Part I: a rationale, some criteria, the conventional tests. Journal of Speech and Hearing Disorders, 45, 431-444 (1980a).

Locke, J., The inference of speech perception in the phonologically disordered child. Part II: some clinically novel procedures, their use, some findings. Journal of Speech and Hearing Disorders, 45, 445-468 (1980b).

Marsh, D., Auditory figure-ground ability in children. The American Journal of Occupational Therapy, 27, 218-255 (1973).

Martin, E., Pickett, J., Sensorineural hearing loss and the upward spread of masking. Journal of Speech and Hearing Research, 13, 426-437 (1970).

Mase, D., Etiology of Articulatory Speech Defects, New York: Columbia University Bureau of Publication, Contributions to Education, 921 (1946).

Mattingly, J., Liberman, A., Syrdal, A., Hawles, T., Discrimination in speech and nonspeech modes. Cognitive Psychology, 2, 131-157 (1971).

McDonald, E., A Deep Test of Articulation, Picture Form, Pittsburg: Stanwix House (1964).

Mecham, M., Measurement of verbal listening accuracy in children. Paper presented at the annual convention of the American Speech and Hearing Association in Chicago, 1969; also published in the Journal of Learning Disabilities, 4, 257-259 (1971).

Miller, C., Morse, P., The "heart" of categorical speech discrimination in young infants. Journal of Speech and Hearing Research, 19, 578-589 (1976).

Miller, C., Morse, P., Dorman, M., Cardiac indices of infant speech perception. Quarterly Journal of Experimental Psychology, 29, 533-545 (1977).

Miller, G., Heise, G., Lichten, W., The Intelligibility of speech as a function of the context of test materials. Journal of Experimental Psychology, 41, 329-335 (1951).

Miller, G., Nicely, P., Analysis of perceptual confusions among some English consonants. Journal of the Acoustical Society of America, 27, 338-352 (1955).

Miller, J., Effects of noise on people. Journal of the Acoustical Society of America, 56, 729-764 (1974).

Miller, J., Noise and children: a review of the literature. Journal of the Acoustical Society of America, 58, 767-780 (1975).

Moffitt, A. R., Consonant cue perception by twenty twenty-four-week-old infants. Child Development, 42, 717-731 (1971).

Morency, A., Auditory modality and reading: research and practice, in Helen K. Smith (ed.), Perception and Reading, Newark, Delaware: International Reading Association (1968).

Morris, W. (ed.), The American Heritage Dictionary of the English Language, Boston: Houghton-Mifflin Company (1971).

Morse, P., The discrimination of speech and nonspeech stimuli in early infancy. Journal of Experimental Child Psychology, 14, 477-492 (1972).

Morse, P., Infant speech perception: origins, processes and alpha centauri, in Minifie, F., and Lloyd, L. (eds.), Communicative and Cognitive Abilities - Early Behavioral Assessment, Baltimore: University Park Press (1978).

Muma, J., Language Handbook Concepts, Assessment, Intervention, Englewood Cliffs, N.J.: Prentice Hall (1978).

Myklebust, H., Learning disabilities: definition and overview, in Progress in Learning Disorders I, New York: Grune and Stratton (1968).

Newby, H., Audiology, New York: Appleton-Century-Crofts (1972).

Norlund, B., Fritzell, B., The influence of azimuth on speech signals. Acta Otolaryngologica, 56, 632-642 (1963).

Owens, E., Intelligibility of words varying in familiarity. Journal of Speech and Hearing Research, 4, 113-129 (1961).

Oyer, H., Kapur, Y., Deal, L., Hearing disorders in the aging: effects on communication, in Oyer, H. and Oyer, E. (eds.), Aging and Communication, Baltimore: University Park Press (1976).

Oyer, H. Doudna, M., Word familiarity as a factor in testing discrimination of hard of hearing subjects. Archives of Otolaryngology, 72, 351-355 (1960).

Pederson, L., Some structural differences in the speech of Chicago Negroes, in Roger W. Shuy (ed.), Social Dialects and Language Learning, Champagne, Illinois: National Council of Teachers of English, Proceedings of the Bloomington, Indiana Conference, 28-61 (1964).

Perozzi, J., Kunze, L., Relationship between speech sound discrimination skills and language abilities of kindergarten children. Journal of Speech and Hearing Research, 14, 382-390 (1971).

Person, R., Essentials of Mathematics (3rd ed.). New York: John Wiley and Sons (1973).

Pollack, I., Rubenstein, H., Decker, L., Intelligibility of known and unknown message sets. Journal of the Acoustical Society of America, 31, 273-279 (1959).

Powers, M., Functional disorders of articulation - symptomatology and etiology, in Travis, L. (ed.), Handbook of Speech Pathology and Audiology, New York: Appleton-Century-Crofts (1971).

Prins, D., Relations among specific articulatory deviations and responses to a clinical measure of sound discrimination ability. Journal of Speech and Hearing Disorders, 28, 382-388 (1963).

Pronovost, W., Dumbleton, C., The Boston University Speech Sound Discrimination Picture Test, Boston: Boston University School of Education (1953).

Rechner, J., Wilson, B., Relation of speech sound discrimination and selected language skills. Journal of Communication Disabilities, 26-30 (1967).

Ringel, R., House, A., Burke, K., Dolinsky, J., Scott, C., Some relations between orosensory discrimination and articulatory aspects of speech production. Journal of Speech and Hearing Disorders, 35, 3-11 (1970).

Rintelmann, W., Schumaier, D., Burchfield, S., Influence of test form on speech discrimination scores of normal listeners. Journal of Auditory Research Supplement 2, 1-30 (1974).

Rintelmann, W., Schumaier, D., Jetty, A., List equivalency and test-retest reliability for normal listeners on Northwestern University Auditory Test No. 6, East Lansing, Michigan: Michigan State University Press (1975).

Ross, M., Lerman, J., Word Intelligibility by Picture Identification, Pittsburg: Stanwix House, Inc. (1971).

Sabatino, D., The construction and assessment of an experimental test of auditory perception. Exceptional Child, 35, 728-737 (1968).

Sabatino, D., Identification of neurologically impaired children through a test of auditory perception. Journal of Consulting Clinical Psychology, 33, 184-188 (1969).

Sanders, D., Aural Rehabilitation, Englewood Cliffs, N.J.: Prentice-Hall (1971).

Sanderson-Leepa, M., Rintelmann, W., Articulation functions and test-retest performance of normal hearing children on three speech discrimination tests: WIPI, PBK-50 and Northwestern University Auditory Test No. 6. Journal of Speech and Hearing Disorders, 41, 503-519 (1976).

Scheifelbusch, R., Lindsey, M., A new test of sound discrimination. Journal of Speech and Hearing Disorders, 23, 153-159 (1958).

Schlanger, B., Galanowsky, G., Auditory discrimination tasks performed by mentally retarded and normal children. Journal of Speech and Hearing Research, 9, 434-440 (1966).

Schultz, M., Schubert, E., A multiple choice discrimination test (MCDT). Laryngoscope, 79, 382-399 (1969).

Schwartz, A., Goldstein, R., Variables influencing performance on speech sound discrimination tests. Journal of Speech and Hearing Research, 17, 25-32 (1974).

Siegenthaler, B., Haspiel, G., Development of two standardized measures of hearing for speech by children. Cooperative Research Program, Project #2372, United States Office of Education (1966).

Silverman, S., Hirsch, I., Problems related to the use of speech in clinical audiometry. Annals of Otorhinolaryngology, 64, 1234-1244 (1955).

Siqueland, E., DeLucia, C., Visual reinforcement of nonnutritive sucking in human infants. Science, 165, 1144-1146 (1969).

Speaks, C., Jerger, J., Method for measurement of speech identification. Journal of Speech and Hearing Research, 8, 185-194 (1965).

Speaks, C., Performance-intensity characteristics of selected verbal materials. Journal of Speech and Hearing Research, 10, 344-347 (1967b).

Stevens, K., House, A., An acoustical theory of vowel production and some of its implications. Journal of Speech and Hearing Research, 4, 303-320 (1961).

Stevens, K., Klatt, D., Role of formant transitions in the voice-voiceless distinction for stops. Journal of the Acoustical Society of America, 55, 653-659 (1974).

Streeter, L., Language perception of two-month-old infant shows effects of both innate mechanisms and experience. Nature, 259, 39-41 (1976).

Swoboda, P., Morse, P., Leavitt, L., Continuous vowel discrimination in normal and at risk infants. Child Development, 47, 459-465 (1976).

Tarnoczy, T., Resonance data concerning nasals, laterals and trills. Word, 4, 71-77 (1948).

Templin, M., A study of sound discrimination ability of elementary school pupils. Journal of Speech Disorders, 8, 127-132 (1943).

Templin, M., Certain Language Skills in Children, Minneapolis: University of Minnesota Press (1957).

Thompson, B., A longitudinal study of auditory discrimination. Journal of Educational Research, 56, 376-378 (1963).

Thorndike, E., Lorge, I., The Teachers Word Book of 30,000 Words, New York: Bureau of Publications, Teachers College, Columbia University (1944).

Thornton, A., Raffin, M., Speech discrimination scores modeled as a binomial variable. Journal of Speech and Hearing Research, 21, 507-518 (1978).

Thornton, A., Raffin, M., Confidence levels for differences between speech discrimination scores - a research note. Journal of Speech and Hearing Research, 23, 5-18 (1980).

Till, J., Infants' discrimination of speech and nonspeech stimuli. Unpublished doctoral dissertation, University of Iowa, Iowa City (1976).

Tillman, T., Carhart, R., Wilber, L., A test for speech discrimination composed of NCN monosyllabic words: Northwestern University Auditory Test No. 6, Technical Report SAM-TR-66-65, USAF School of Aerospace Medicine, Aerospace Medical Division (AFSC), Brooks Air Force Base, Texas (1966).

Tillman, T., Olsen, W., Speech audiometry, in Jerger, J. (ed.), Modern Developments in Audiology (Second Edition), New York: Academic Press (1974).

Travis, L., Rasmus, B., The speech sound discrimination ability of cases with functional disorders of articulation. Quarterly Journal of Speech, 17, 217-226 (1931).

Trehub, S., Infants' sensitivity to vowel and tonal contrasts. Developmental Psychology, 9, 81-96 (1973b).

Trehub, S., Rabinovitch, M., Auditory-linguistic sensitivity in early infancy. Developmental Psychology, 6, 74-77 (1972).

Utley, J., Whats Its Name, Urbana, Illinois: University of Illinois Press (1951).

Wachs, T., Uzgiris, I., Hunt, J., Cognitive development in infants of different age levels and from different environmental backgrounds. Merrill-Palmer Quarterly Journal of Behavior Development, 17, 288-317 (1971).

Weinberg, G., Schumaker, J., Statistics: An Intuitive Approach. Belmont, Ca.: Wadsworth (1969).

Weiner, P., Auditory discrimination and articulation. Journal of Speech and Hearing Disorders, 32, 19-28 (1967).

Wepman, J., Auditory discrimination: its role in language comprehension, formulation and use. The Pediatric Clinics of North America, 15, 721-727 (1968).

Wepman, J., Auditory Discrimination Test (Manual of Directions), Chicago: Language Research Associates (1958) (Rev. 1973).

White, S., The effect of response availability on the speech discrimination scale. Unpublished doctoral dissertation, Wayne State University, Detroit, Michigan (1974).

Willeford, J., Central auditory function in children with learning disabilities. Audiology and Hearing Education, 2, 12-20 (1976).

Wilson, R., Antablin, J., A picture identification task as an estimate of word recognition performance of nonverbal adults. Journal of Speech and Hearing Disorders, 45, 223-238 (1980).

Winer, B., Statistical Principals in Experimental Design. New York: McGraw-Hill (1971).

Winitz, H., Articulatory Acquisition and Behavior, New York: Appleton-Century-Crofts (1969).

Winitz, H., Bellrose, B., Sound discrimination as a function of pre-training conditions. Journal of Speech and Hearing Research, 5, 340-348 (1962).

Winitz, H., Bellrose, B., Effects of pre-training on sound discrimination learning. Journal of Speech and Hearing Research, 6, 171-180 (1963).

Wood, G., Fundamentals of Psychological Research, Boston: Little, Brown and Company (1974).

Wood, K., Terminology and nomenclature, in Travis, L. (ed.), Handbook of Speech Pathology and Audiology, New York: Appleton-Century-Crofts (1957).

Woolf, G., Pilberg, R., A comparison of three tests of auditory discrimination and their relationship to performance on a Deep Test of Articulation. Journal of Communication Disorders, 3, 239-249 (1971).

APPENDICES

APPENDIX A

THE GOLDMAN-FRISTOE-WOODCOCK TEST OF AUDITORY DISCRIMINATION, QUIET SUBTEST

CASH	CAP
WAKE	BEAR
DIG	LAKE
ME	WE
FAIR	SIGN
CATCH	COAL
TACK	MAIL
RAKE	PACK
KNEE	SAIL
JACK	BEE
BIG	SHACK
VINE	TEA
NIGHT	MAKE
CONE	BACK
PAIL	HAIR

APPENDIX B

THE WORD INTELLIGIBILITY BY PICTURE IDENTIFICATION SPEECH DISCRIMINATION TEST, LIST I

SCHOOL	WING
BALL	MOUSE
SMOKE	SHIRT
FLOOR	GUN
FOX	BUS
HAT	TRAIN
PAN	ARM
BREAD	CHICK
NECK	CRIB
STAIR	WHEEL
EYE	STRAW
KNEE	PAIL
STREET	

APPENDIX C

PARENTAL CONSENT AND RELEASE FORM

I hereby consent to allow my child to participate as a subject in the speech discrimination research project to be carried out at the Northern Arizona University Speech and Hearing Center.

I understand that I am giving this consent voluntarily, with no expectation of payment. I also understand that I am under no obligation to participate, and that I may withdraw from participation at any time with no penalty.

I am aware that, once my child is accepted as a subject in this project, any data obtained with him or her will be used for research purposes only, on a strictly anonymous basis. Any findings specifically related to my child will be made known to me, if I desire. Original copies of all test materials used will remain the property of Mr. William Culbertson, the chief researcher involved in this project.

____
Signature

____
Date

APPENDIX D

RAW DATA OBTAINED WITH FIVE-YEAR-OLD SUBJECTS IN THE QUIET LISTENING CONDITION (REPEATED PRESENTATIONS)

		Number of Correct Responses on WIPI/GFW at each Sensation Level				
Subject	SRT (dB HTL)	8 dB	16 dB	24 dB	32 dB	40 dB
K.B.	8	17/27	25/30	25/30	25/30	25/30
T.H.	14	18/27	20/29	25/30	25/30	25/30
L.P.	20	17/26	28/27	24/29	25/30	25/30
N.M.	18	15/24	21/28	22/29	24/26	25/28
J.B.	20	9/16	16/22	20/23	21/25	21/29
A.H.	18	20/25	23/28	25/30	25/30	25/30
W.J.	24	9/26	25/30	25/30	25/30	25/30
H.S.	26	21/26	27/30	25/30	25/30	25/30
R.S.	20	17/22	23/26	25/27	24/27	24/27
S.W.	18	11/19	23/25	25/30	25/30	25/30

APPENDIX E

RAW DATA OBTAINED WITH FIVE-YEAR-OLD SUBJECTS IN THE NOISE LISTENING CONDITION (REPEATED PRESENTATIONS)

		Number of Correct Responses on WIPI/GFW at Each Sensation Level				
Subject	SRT (dB HTL)	8 dB	16 dB	24 dB	32 dB	40 dB
D.W.	12	13/16	23/31	24/22	23/27	25/26
J.S.	16	18/17	19/23	18/25	18/25	20/25
A.P.	16	14/21	18/19	20/22	22/22	21/20
D.O.	14	11//4	19/19	20/22	21/2/	21/22
L.O.	16	18/24	22/28	24/29	24/30	24/30
J.C.	16	12/23	22/29	23/30	25/30	25/30
J.S.	12	20/24	23/30	24/30	23/28	24/30
T.G.	26	15/24	22/26	24/26	23/30	23/30
V.V.	14	14/21	19/23	17/22	21/21	20/24
L.H.	18	9/15	18/19	22/22	24/21	22/23

APPENDIX F

RAW DATA OBTAINED WITH TEN-YEAR-OLD SUBJECTS IN THE QUIET LISTENING CONDITION (REPEATED PRESENTATIONS)

Subject	SRT (dB HTL)	Number of Correct Responses on WIPI/GFW at Each Sensation Level				
		8 dB	16 dB	24 dB	32 dB	40 dB
T.O.	16	16/22	24/26	25/28	25/30	25/30
D.K.	26	21/26	25/30	25/29	25/30	25/30
A.R.	12	18/21	23/27	25/29	25/30	25/30
R.S.	14	20/24	24/27	25/28	25/30	25/30
P.H.	20	22/30	25/30	25/30	25/30	25/30
L.B.	12	12/30	24/30	25/30	25/30	25/30
B.D.	12	21/24	23/30	25/30	25/30	25/30
J.K.	8	5/24	15/27	25/30	25/30	25/30
R.S.	16	18/24	24/30	24/29	25/30	25/30
T.F.	16	18/21	25/28	25/29	25/29	25/30

APPENDIX G

RAW DATA OBTAINED WITH TEN-YEAR-OLD SUBJECTS IN THE NOISE LISTENING CONDITION (REPEATED PRESENTATIONS)

		Number of Correct Responses on WIPI/GFW at Each Sensation Level				
Subject	SRT (dB HTL)	8 dB	16 dB	24 dB	32 dB	40 dB
A.M.	8	14/20	21/24	22/29	22/28	22/27
L.S.	8	14/18	20/25	23/26	24/28	25/28
C.S.	8	13/18	21/26	22/27	23/26	23/27
D.G.	14	16/19	22/26	25/30	25/30	25/30
T.C.	12	25/22	21/28	23/28	23/28	23/30
E.B.	12	19/27	24/30	24/30	24/30	24/30
A.W.	12	16/28	20/30	24/30	23/30	23/30
W.M.	12	19/21	25/25	25/30	25/30	25/30
T.W.	14	16/21	22/30	22/30	25/30	25/30
L.H.	12	20/26	23/30	24/30	24/30	25/30

APPENDIX H

RAW DATA OBTAINED WITH THE FIVE-YEAR-OLD SUBJECTS IN BOTH LISTENING CONDITIONS (SINGLE PRESENTATION)

Quiet Listening Condition			Noise Listening Condition		
Subject	SRT (dB HTL)	No. Correct (WIPI/GFW)	Subject	SRT (dB HTL)	No. Correct (WIPI/GFW)
S.M.	18	23/27	J.L.	14	18/25
M.K.	14	24/27	C.M.	18	19/23
B.B.	18	25/27	J.C.	32	19/20
R.O.	20	22/24	M.J.	20	21/25
M.S.	18	22/25	J.S.	18	22/27
S.M.	20	24/28	S.L.	24	13/18
K.K.	20	20/25	M.L.	14	21/25
A.E.	10	21/28	K.R.	20	20/23
N.R.	18	24/28	J.W.	16	19/18
B.S.	22	23/22	C.S.	18	19/28

APPENDIX I

RAW DATA OBTAINED WITH THE TEN-YEAR-OLD SUBJECTS IN BOTH LISTENING CONDITIONS (SINGLE PRESENTATION)

Quiet Listening Condition			Noise Listening Condition		
Subject	SRT (dB HTL)	No. Correct (WIPI/GFW)	Subject	SRT (dB HTL)	No. Correct (WIPI/GFW)
A.R.	12	23/29	A.B.	14	22/29
B.V.	14	25/30	S.S.	10	23/24
J.W.	8	23/28	N.C.	14	23/25
J.N.	12	25/29	C.J.	16	23/28
B.H.	12	25/29	L.L.	16	22/26
L.L.	10	22/29	S.L.	12	21/26
B.K.	14	23/28	J.R.	14	22/26
S.M.	14	25/26	J.N.	14	23/27
C.S.	14	25/27	M.S.	14	20/25
R.B.	10	23/27	M.M.	10	21/27

APPENDIX J

DATA COLLECTION FORM

NAME: ______________________ AGE: ______________

LISTENING CONDITION: ______________ FIRST TEST: ______________

NUMBER: ____________ SRT: ______________________

WIPI	SL (dB)	#ERRORS	SCORE (%)		
				School	Wing
				Ball	Mouse
	8			Smoke	Shirt
				Floor	Gun
	16			Fox	Bus
				Hat	Train
	24			Pan	Arm
				Bread	Chick
	32			Neck	Crib
				Stair	Wheel
	40			Eye	Straw
				Knee	Pail
				Street	

GFW	SL (dB)	#ERRORS	SCORE (%)		
				Cash	Cap
				Wake	Bear
	8			Dig	Lake
				Me	We
	16			Fair	Sign
				Catch	Coal
	24			Tack	Mail
				Rake	Pack
	32			Knee	Sail
				Jack	Bee
	40			Big	Shack
				Vine	Tea
				Night	Make
				Cone	Back
				Pail	Hair