Comparison of word-unit and nonsense-unit speech sound discrimination scores of preschool children

Mary Pat LaForest

The University of Montana

Follow this and additional works at: https://scholarworks.umt.edu/etd

Let us know how access to this document benefits you.

Recommended Citation

LaForest, Mary Pat, "Comparison of word-unit and nonsense-unit speech sound discrimination scores of preschool children" (1973). Graduate Student Theses, Dissertations, & Professional Papers. 8041. https://scholarworks.umt.edu/etd/8041

This Thesis is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Graduate Student Theses, Dissertations, & Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.
COMPARISON OF WORD-UNIT AND NONSENSE-UNIT SPEECH SOUND DISCRIMINATION SCORES OF PRESCHOOL CHILDREN

By

Mary Pat Sullivan LaForest

B.A. University of Montana, 1971

Presented in partial fulfillment of the requirements for the degree of

Master of Arts

UNIVERSITY OF MONTANA

1973

Approved by:

[Signatures]

Chairman, Board of Examiners

Dean, Graduate School

Date

Aug 14, 1973
ACKNOWLEDGEMENTS

The writer wishes to thank her thesis committee, Dr. Kellogg Lyndes, Dr. Laurence H. Berger, and Mrs. Sara McClain, for their advice and comments in the preparation of this thesis.

Gratitude is also expressed to the University of Montana Nursery School and the Playmate Nursery for their cooperation and assistance in carrying out this study.

Finally, the writer wishes to express her sincere appreciation to Dr. Evan P. Jordan for his patient guidance and encouragement throughout this study.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II. PROCEDURE</td>
<td>14</td>
</tr>
<tr>
<td>III. RESULTS AND DISCUSSION</td>
<td>26</td>
</tr>
<tr>
<td>IV. SUMMARY AND CONCLUSIONS</td>
<td>45</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>47</td>
</tr>
<tr>
<td>APPENDIX</td>
<td>50</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Cut-off Scores, Mean Scores and Standard Deviations on the Templin-Darley Screening Test of Articulation by Age for Boys and Girls Combined</td>
</tr>
<tr>
<td>2. Summary of the Analysis of Variance of Errors Per Fifth on the Speech Sound Discrimination Test</td>
</tr>
<tr>
<td>3. Summary of the Analysis of Variance of Scores on an Experimental Speech Sound Discrimination Test</td>
</tr>
<tr>
<td>4. Mean Number of Errors for Meaning, Position and Sex on the Experimental Test of Discrimination</td>
</tr>
<tr>
<td>5. Correlations Between Scores for the Variables of Meaning and Position</td>
</tr>
<tr>
<td>6. Summary of Analysis of Variance of &quot;Same&quot; Pair Errors</td>
</tr>
<tr>
<td>7. Percentages for Total Number of Errors on Meaningful and Nonmeaningful Test Stimuli</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
</tr>
<tr>
<td>28</td>
</tr>
<tr>
<td>30</td>
</tr>
<tr>
<td>31</td>
</tr>
<tr>
<td>33</td>
</tr>
<tr>
<td>42</td>
</tr>
<tr>
<td>42</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

Auditory discrimination is most frequently defined in terms of speech sound discrimination, which, in turn, has been defined as the ability to recognize or distinguish between phonemes that are closely related or highly similar in phonetic feature terms (Weiner, 1967, and Wepman, 1960). It is in this sense that the term speech sound discrimination will be used in this paper. The phenomenon of speech sound discrimination has long been of interest to speech pathologists and audiologists. It has been explored from many auditory aspects, such as the nature of the auditory signal and the psychologic set of the listener (Wood, 1971).

A vital consideration in the exploration of any phenomenon is the nature of the measuring instrument used in the attempt to gain increased information. The phenomenon is never known in and of itself, but only through the lens, both clarifying and distorting, of the measuring devise. The investigations of auditory discrimination are in no sense exceptions of this truism (Weiner, 1967).

A variety of tests have been constructed for use in investigation of auditory discrimination, tests which vary in both form and content. This study was concerned with the effect of the meaningfulness of the auditory stimuli used in phonemic discrimination tests.

Prior to any discussion of discrimination, the terms, phoneme, phonemic and phonetic need to be defined. "Phoneme" is a technical term for "a range of sounds that the speakers of a given language
perceive as functionally the same and discriminate from other ranges of sounds" (Carroll, 1964). The term "phonetic" pertains to sound production; it refers to the descriptive structure of sounds. "Phonemics" refers to the study of the sound systems of a language.

Phonemes are not considered meaningful in themselves, but make up and provide the critical bases for differentiating among the forms of language which are meaningful or grammatically functional. Presumably a child learns the phonemes of his language through a process of gradual differentiation. It has been hypothesized that the child is reinforced for making the phonemic judgments necessary for his language and receives no reinforcement for incorrect phonemic discrimination. He learns to recognize a group or range of sounds as one phoneme and to discriminate between the range of sounds (allophones) and one phoneme and the range (allophones) of another phoneme. In other words, a child learns to recognize the sound differences which make a difference for his language.

Phonetic discrimination refers to discrimination between sounds which are classified as belonging to the same phoneme. There are few, if any, practical tests of phonetic discrimination. Liberman et al. (1957) investigated the relation between phonemic labeling and discrimination in one language and within one group of phonemes (the latter, a phonetic discrimination task). A synthesizer was used to generate speech-like sounds for this experiment. Subjects were asked to first identify sounds, presented singly, thus dividing the stimuli into three phonemic categories: b-d-g. In the second part, the subjects were asked to discriminate within each sound or discriminate allophones.
of the identified phoneme categories; this was a phonetic discrimination task. Liberman et al. found better discrimination across phoneme boundaries (between phonemes) than within phoneme boundaries (phonetic discrimination).

Phonemic discrimination refers to differentiating the allophones of one phoneme from all of the allophones of another phoneme. One of the most common ways to identify and discriminate phonemes of a language is to use pairs of forms (syllables or words for example) which are phonemically the same except for one sound, such as "pin" and "bin". These are called "minimal pairs" (Carroll, 1964, and Dale, 1972). An example of a phonemic discrimination test is the Wepman Auditory Discrimination Test. In this test, the child is asked to listen to either minimal or identical word pairs and to indicate whether each pair of words is the "same" two words or two "different" words. Theoretically, the child's task is to discriminate one phoneme and its allophones from another phoneme and all of its allophones.

Berry (1969) theorized that the discrimination tests designated as tests of phonemic discrimination, such as the Wepman, are in reality tests of free morphemic discrimination. (A free morpheme is defined as a unit of speech that is recurrent, meaningful and can stand alone; most of these are words.) She cited evidence to show that a phoneme of a language need not be discriminable before its symbolic significance in the morpheme can be appreciated and that the morpheme is the smallest discriminable acoustic unit. From this viewpoint, it would appear that a child learns to discriminate the phonemes of his language through the process of successive discriminations of the morphemes of the language.
It is this author's opinion that it is a theoretical question whether a child makes a phonemic or morphemic discrimination and as such is not measureable by presently available instruments. To date, there is no available normative data on the effect of the meaningfulness of test stimuli on a measurement of a child's speech sound discrimination. A comparison of meaningful versus nonmeaningful auditory stimuli seems warranted in terms of normative data and future discrimination testing.

Review of the Literature

For the purpose of the procedures of this study, a review of the major research conclusions of phonemic discrimination tasks and the instruments used to obtain such data seems warranted. Both Wepman (1960) and Weiner (1967) summarized the major research conclusions, which included the following:

(1) There is evidence that the more nearly alike two phonemes are in phonetic structure, the more likely they are to be misdiscriminated (Liberman, 1957).

(2) Individuals differ in their ability to discriminate among sounds (Templin, 1957).

(3) Auditory discrimination shows a developmental progression and frequently matures as late as the end of the child's eighth year (Templin, 1957).

(4) There is a positive relationship between slow development of auditory discrimination and articulation defects (Kronvall and Diehl, 1954; Cohen and Diehl, 1963; Schiefelbusch and Lindsey, 1958; Sherman and Geith, 1967).

Speech sound discrimination has been investigated using a number of different kinds of test instruments. Of importance to this study is the orally-presented stimuli which are used in the discrimination tasks. Many of these instruments utilize meaningful test stimuli, such as
words and/or pictures, and some have used nonmeaningful material, such as nonsense syllables.

One of the earliest tests of sound discrimination was constructed by Travis and Rasmus in 1931. This test has been considered the most extensive ever used in discrimination studies (Wepman, 1960). Every consonant and vowel in the English language was compared with every other consonant or vowel and with itself, resulting in 366 minimal or identical pairs of nonsense syllables. The subject was to make a "same" or "different" judgment after hearing each pair. The overwhelming task of making 366 discrimination comparisons was quickly recognized. In 1943, Templin presented a discrimination test for children, age six to eight, which was based on the Travis-Rasmus model. The test was composed of 70 pairs of nonsense syllables; the pairs were either identical syllable pairs or minimal pairs differing by one phoneme. Again, the test called for a "same" or "different" judgment by the child. Templin (1957) pointed out that these tests demand that a subject understand the concepts "same" and "different", and that using nonsense syllables increased the difficulty of the task.

In the Wepman Auditory Discrimination Test (1958) the method of presenting sound pairs (which are identical or differ by only one sound) is retained from the Travis-Rasmus model, but 40 word pairs rather than nonsense syllables are used as the test stimuli. Weiner (1967) notes that the "abstract" character of the test is reduced, as is the difficulty of the test, since a difference in meaning rather than merely in sound value is created by the use of real words. The Travis-Rasmus, Templin and Wepman tests employed the same form in their tests but used
different test stimuli. There is no investigation of the possible effects that the meaningfulness of the test stimuli had on the resulting discrimination scores from these measures.

In the literature, nonsense syllables are generally considered more difficult to discriminate than words (Goetzinger, 1972). An advantage of nonsense syllables in testing speech discrimination is their lack of meaning and thus, the listener's vocabulary is not a variable. As test material, Goetzinger considers nonsense syllables an excellent measure of "pure" speech sound discrimination. The lack of meaning, in turn, can also be a disadvantage since the listener does not normally need to identify meaningless speech stimuli. Carhart (1965) noted that nonsense syllables are confusing to many subjects. It has also been noted that nonmeaningful material does not possess the motivational value of meaningful material. At this time the differential effects of using nonsense syllables in testing discrimination as compared to using words is not known.

Speech discrimination tests are generally included in audiological assessment. A listener's ability to hear words is often reported by a percentage of the words heard or is referred to as an articulation score. The effect of using different test stimuli has been investigated, but in terms of this articulation testing. Articulation tests were developed in which an announcer read a list of syllables, words or sentences to a group of adult listeners; the listeners recorded the items as they were read. When nonsense syllables were used as test stimuli, trained listeners recorded the nonsense items phonetically. The percentage of items correctly recorded by these listeners was
called the articulation score. The basic concern with this type of test was (1) the increase in the number of words that became intelligible as speech became louder, (2) what was the maximum percentage of words, syllables, etc. intelligible at the most favorable intensity, and (3) what was the intensity level of speech which was maximally intelligible to a given subject (Davis and Silverman, 1970, and Licklider and Miller, 1958). Licklider and Miller (1957) reviewed the articulation test methods and findings of Fletcher (1929), Egan (1943) and Hudgens et al. (1947). Results of the articulation testing were reported by plotting the percentage of items heard correctly against the relative intensity of the speech in decibels. Overall results indicated the following in order of increasing percentage of correct scores: nonsense syllables, monosyllabic words, polysyllabic words and sentences. With adults, the use of nonsense syllables yielded somewhat lower percentages of correct responses at any given intensity level than did words. However, these investigations did not include children, nor was the effect of the test stimuli on discrimination ability at any one intensity level investigated.

Elenbogen and Thompson (1972) compared the discrimination ability of lower class children using the Wepman Test and a test composed of nonsense words which were created by exchanging the phonemes of the word pairs on the Wepman. For example, "tub-tug" on the Wepman became "teb-teg" on the distorted or nonsense test. Kindergartners of middle class backgrounds scored higher on the Wepman than on the distorted Wepman. Kindergartners of lower socioeconomic status had nearly identical scores on both tests. It appeared that it was more difficult for
kindergartners of middle class backgrounds to discriminate between phonemes when meaning was removed from test items, but this did not appear to be true for kindergartners of lower socioeconomic status. It was concluded that a vocabulary factor was measured along with speech sound discrimination in the Wepman.

Elenbogen and Thompson did not measure or control the articulation skills or hearing acuity of the subjects used in their investigation. Previous research has indicated that both articulation skills and hearing acuity are related to speech sound discrimination ability. Elenbogen and Thompson did not attempt to use word and nonsense syllable items that were as phonemically alike as possible. For example, in exchanging phonemes to create a nonsense syllable item, vowels were changed (tub-tug became teb-teg), test items differed by more than one phoneme (leg-led became mĩg-mĩd), and manner or place of production of phonemes was not matched (lack-lack became chak-chak). The present study was concerned with the effect of meaningfulness on speech sound discrimination scores when word and nonsense syllable items were matched closely with respect to phonemic content and only subjects with normal hearing and articulation skills were used.

In a number of discrimination tests, the subject has been asked to respond to stimuli in two sense modalities—the auditory and the visual. In the Schiefelbusch and Lindsey (1958) test, pictures were presented to the subject and he must tell whether or not the words represented by the pictures sound alike. Discrimination of initial, final and rhyming sounds was required. In a number of auditory-visual discrimination tests, a subject has been asked to point to a picture named by
the tester. Templin's Picture Discrimination Test (1957) is an example. A number of authors have indicated that the use of pictures is particularly helpful in maintaining interest with young children. Most tests of this kind either include a pretraining session to familiarize the subject with the pictures so that vocabulary (i.e., meaningfulness) is not a factor or construct the test with words that have been found to be familiar for the particular age group being tested. However, in Templin's 1957 study, the correlations between the children's ability to name the pictures and to make the discriminations were extremely high at each age level (.91 to .96). Of course, such a task is not practical with nonsense syllables; thus, these types of tests (picture naming) would not be appropriate for comparing the effects of words and nonsense syllables on discrimination.

A subject's own articulation errors comprise part of the content of several discrimination tasks. Such tests could be labeled "deep tests" of discrimination since one sound is tested in a variety of sound contexts. Aungst and Frick (1964) investigated the relation between the production and discrimination of the /r/ phoneme. Discrimination for the /r/ phoneme was tested in the following ways: (1) the child's immediate evaluation of his own response, (2) a delayed judgment in which the child evaluated his own response when heard from a tape recording, and (3) the child's evaluation of his own response compared to the experimenter's response. The child's task was to judge the correctness of the /r/ productions. Farquhar's test (1961) required children to clap hands when they heard the examiner utter the correct form of their misarticulated sound in a series of trials. Anderson
investigated discrimination of /s/ in various contexts by imitating the child's misarticulation of that sound as closely as possible in one of three utterances of a word. The child's task was to indicate the incorrect production. All of these "deep tests" used words for test stimuli. One could argue that when an incorrect sound was used, the resulting test word was not a standard English word and thus relatively nonmeaningful. However, again the possible effect of using nonsense syllables versus words in discrimination testing has not been systematically investigated.

The terms meaningful and nonmeaningful have been used throughout this discussion. The concept of "meaningfulness" is a difficult one to define. Carroll (1964) states that a meaning response associated with a word will expand in connotative meaning when the word is experienced in a wide variety of contexts; the richness of the connotations may be called "meaningfulness". C. E. Noble (1961) measured the meaningfulness of words based on the rate at which subjects gave verbal associations to a word. It should be noted that words of rare or limited use and nonsense syllables have been found to have low degrees of meaningfulness. Noble used trigrams (groups of three letters) and asked subjects to rate meaningfulness on a 0 to 7 scale ("0" representing no meaning and "7" most meaningful). However, his trigrams were based on the letters of the alphabet and not on phonetic transcriptions. There is no available data rating the meaningfulness of C-V-C syllables presented orally. For the purpose of this study: (1) meaningful stimuli will refer to words which occur frequently in American English and therefore will have a relatively high probability of representing
meaningful units, and (2) nonmeaningful stimuli will refer to nonsense syllables devoid of any denotative meaning and therefore presumed to have relatively little connotative meaning.

Statement of the Problem

The primary purpose of this study was to investigate the possible effects of meaningfulness of test stimuli on speech sound discrimination. An experimental speech sound discrimination test was constructed in which half of the test items were meaningful test stimuli (words) and the other half were nonmeaningful test stimuli (nonsense syllables). Error scores obtained on the meaningful and nonmeaningful test items were compared. The primary hypothesis tested was that there would be a difference in scores between the word items and the nonsense syllable items (with fewer errors on word items), when both were presented orally to a group of children.

In the construction of this experimental discrimination test, 75 percent of the total test items were minimal paired items, differing by one consonant phoneme in either prevocalic or postvocalic word positions. There seems to be some controversy in the literature concerning discrimination performance in the prevocalic and postvocalic positions. Templin (1943), in a study of speech sound discrimination ability of elementary school students, grades second through sixth, found that sound contrasts in the medial or final (postvocalic) position were more difficult to discriminate than initial (prevocalic) position sound contrasts. Templin used two 100-item nonsense syllable discrimination tests which were identical except for the position of the consonant contrast; the first 100 items involved contrasts in the initial or prevocalic position and
the second 100 items included item contrasts in the medial or final (postvocalic) position. Kamil and Rudegeair (1972) matched consonant sounds in both initial (prevocalic) and final (postvocalic) positions in nonsense syllable items and administered similar forms of a discrimination test on successive days to kindergarten and first grade students. They found that, overall, there was no difference in the performance on prevocalic and postvocalic contrast pairs. However, it is not known whether the position of the sound contrasts had any significant effect in the first administration of Kamil and Rudegeair's discrimination test since these results were not published. In the present study, it was possible to note any differential effect of the sound contrast position and the differential effect of meaningfulness with regard to position. It was predicted that there would be significant difference in error scores with regard to sound contrast position, with final sounds showing more errors of discrimination.

Templin (1957) administered a word-picture discrimination test to children, ages three through five years. It was found that while girls scored slightly higher than did boys from age four on, the sex differences were not significant. Templin (1957) also administered a nonsense syllable discrimination test to children, ages six through eight years. For this entire age range, the mean scores of the sexes were significantly different, with girls scoring higher than boys. In the present study, the differential effect of the sex of the subject was also investigated; it was predicted that girls would make fewer errors than boys, but that the sex differences would not be significant at conventional levels of significance (.05 level).
In summary, the experimental variables evaluated in this study were: (1) meaning, (2) position, and (3) sex. The hypotheses tested were:

(1) That there would be a statistically significant difference in scores between the word items and the nonsense syllable items;

(2) That there would be significant differences in scores between prevocalic and postvocalic sound contrasts; and

(3) That there would be no significant sex differences in the resulting scores.
CHAPTER II

PROCEDURE

Materials

Two lists of 40 pairs of speech sound discrimination test items were constructed. One list was comprised of word pairs and the other nonsense syllable pairs. The items were matched with respect to phonemic content, as described below. The task required that a child listen to a taped reading of a combination of both lists of items and indicate whether each pair of words or syllables was the "same" word read twice or two "different" words.

Both word pair and nonsense syllable pair lists were composed of 10 identical pairs and 30 minimal pairs, word or syllable pairs phonetically the same with the exception of one consonant phoneme. All test items were equated for length and syllable structure by using the word or syllable structure: consonant-vowel-consonant (CVC). The consonant phonemes contrasted in the 30 minimal pairs of words or syllables were consonant phonemes often misarticulated by normal kindergarten children (Snow, 1963). Nineteen consonant sounds plus two glides and two liquids with their most frequent substitution in the prevocalic and postvocalic positions were selected from Snow's data. For this study, each of the selected sounds was contrasted with its usual substitution in both

\(^1\)Two exceptions to the CVC rule were words containing /ts/ and /dz/ as frequent substitutions for /ʃ/ and /ʒ/, respectively.
prevocalic and postvocalic positions. Thus, 15 of the minimal pairs contrasted by the prevocalic consonant and 15 of the minimal pairs contrasted by the postvocalic consonant. The word pairs were matched for frequency of occurrence and familiarity by choosing from the Lorge-Thorndike, Teacher's Word Book of 30,000 Words (1944). The majority of the word pairs were selected from the 1,000 most frequently occurring words; the remainder of the word selection was from the 2,000 to 5,000 most frequently occurring words. Selection of these criteria for the word pairs was an attempt to insure with some degree of probability that the children would be familiar with the test words and, thus, the test items would represent relatively highly meaningful units.

Words and syllables were paired such that a frequently misarticulated phoneme contrasted with its most usual substitute, for example, /f/ and /θ/ ("fought" and "thought" on the word list and /fɔp/ and /θɔp/ on the nonsense syllable list). The contrasted phonemes and the vowel from each paired word item were duplicated in the nonsense syllable pairs, as exemplified above. Whenever possible, the non-contrasting consonant for each nonsense syllable pair was selected from those consonants having the same manner of production as the consonant in the similar word pair. Only two of the non-contrasting consonants were selected on the basis of the same place of production rather than the

---

2 Four sounds and their frequent substitutions (w/r, w/l, ts/y, dz/dʒ) were contrasted in the word position in which one of the pair occurs in English. For example, /ts/ and /ʃ/ contrasted in the postvocalic position in the words, "coats-coach"; /ts/ does not occur prevocally in English.

3 To develop minimal pairs, it was necessary to use one word which did not occur as frequently as to fall among the first 5,000 words.
same manner; this was necessary to create a nonsense syllable rather than another word. In the above example, the final plosive /t/ in "fought" and "thought" was replaced by the plosive /p/, in the nonsense syllable pairs.

When the word pair items and the nonsense syllable pair items were constructed, the total of 80 items were combined to form one experimental speech sound discrimination test. Twenty-five percent of the test items were identical or "same" pairs. Similarly, psychological research has a usual distribution of 25 percent "same" trials, which are often considered "blank" trials. For this study, it was arbitrarily decided that the 25 percent "same" pairs would be distributed so that every fourth item was an identical or "same" pair in the experimental discrimination test. The word and nonsense syllable "same" pairs were alternated so that the fourth item was a word "same" pair and the eighth item was a nonsense syllable "same" pair, etc. The minimal pair items were also alternated with a word pair followed by a nonsense syllable pair. The test was developed into five equal sections or fifths. This was done to provide a way to analyze any fatigue factor which may have occurred and then eliminate a section or fifth where fatigue was demonstrated. Fifths were selected because of mathematical ease of dividing the items. The sound contrasts in both prevocalic and postvocalic positions for the word minimal pairs were randomly assigned to each fifth. The same groups of contrasts for fifths was retained for the nonsense syllable pairs but the order was changed. For example, the contrasts for the word minimal pairs (v/f, θ/s and p/b) were in the third section and the same sound contrasts for the nonsense pairs were in the first section.
Validation Procedure

Since word pairs were used for meaningful stimuli, it was of concern whether the words selected actually represented words to the children. A pilot procedure was conducted to determine whether children recognized the word test items as words. The investigator read the word list along with nonsense syllable items, as "foils", to a group of eight children, the same ages as those included in the study. Each child was asked whether the item he heard was a word or not a word. On the majority of the word items there was 100 percent agreement (all of the children indicated "yes"). Only items which 75 percent of the children (six of eight) indicated were words, were included as test stimuli in the experimental test. It should be noted that it was not necessary to delete any word item on the basis of this validation procedure. It is the experimenter's opinion that this was a crude procedure, but the investigation of the meaningfulness of words or awareness of words by preschool children presented an entirely different study and could not be dealt with in any depth in this study.

Subjects

Twelve girls and 12 boys were selected from nursery school classes in the Missoula area. The subjects ranged in age from 4 years-6 months to 5 years-6 months, with a mean age of 5 years-2 months. This age range was selected for two main reasons: (1) according to Templin (1957), children in this age category are still developing discrimination skills, and (2) a number of investigators have noted the difficulty of teaching the same-different concept to younger children (Templin, 1957).
Screening Procedures

The Templin-Darley Screening Test of Articulation was administered to each subject to insure "average" articulation development, since research has indicated a relationship between slow development of auditory discrimination and articulation defects. The cut-off scores from the Templin-Darley Screening Test (Templin-Darley manual, Table 3, p. 25) were used as a means to eliminate children with inadequate or defective articulation from this experiment. As an upper limit for the articulation screening, it seemed reasonable to include children who scored within one standard deviation above the mean score of their age level for the sexes combined (Templin-Darley manual, Table 2, p. 25). Since the Templin-Darley does not provide a cut-off score or mean score for the age interval 5 years-6 months, these scores were extrapolated from the given scores. The mean scores, cut-off scores and standard deviations on the Templin-Darley Screening Test are provided in Table 1; these scores have been rounded off to the nearest whole number.

Children ages 4 years-6 months to 4 years-9 months were grouped into the 4½ year interval, children ages 4 years-10 months to 5 years-

<table>
<thead>
<tr>
<th>Age</th>
<th>4½</th>
<th>5</th>
<th>5½</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cut-off Score</td>
<td>26</td>
<td>31</td>
<td>33</td>
</tr>
<tr>
<td>Mean Score</td>
<td>36</td>
<td>38</td>
<td>40</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>12</td>
<td>13</td>
<td>12</td>
</tr>
</tbody>
</table>

Table 1. Cut-off Scores, Mean Scores and Standard Deviations on the Templin-Darley Screening Test of Articulation by Age for Boys and Girls Combined.
2 months were grouped into the 5 year interval and children ages 5 years-3 months to 5 years-6 months were grouped into the 5½ year interval. Only children who scored at or above the cut-off score and at or below the upper limit of one standard deviation above the mean for their age level were included in the experiment as subjects.

Socioeconomic status was not considered in this study as a control. Templin (1957) and Weaver et al. (1960) found that children of lower socioeconomic status made more articulation errors than children of upper socioeconomic status. In this study, only children with "average" articulation skills were included without regard to socioeconomic background. It was assumed that by setting articulation skill cut-off scores, any negative influence of socioeconomic status was indirectly eliminated. Research has indicated that children of different socioeconomic backgrounds tend to perform or react differently to a testing situation. Again, any child who passed the concept and task teaching trials was included in the study, regardless of background.

Each subject was given a screening pure tone hearing test in a sound-treated suite to determine normal hearing acuity. All subjects chosen for this study responded at the 15 dB level, bilaterally, for the frequencies 500, 1000, 2000 and 4000 Hz (as described by Davis and Silverman, 1970).

**Instrumentation**

The speech sound discrimination test was recorded on a Sony TC-230 4-track tape recorder. A memorex 1.5 mil x 1200 tape was used for the recording at a tape speed of 7½ in./sec. An adult female read the test items with a one second pause between word or syllable pairs and a ten

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
second pause between paired items. Each pair was preceded by the com-
mand "Listen" and followed by the question, "Alike or not alike?".
The test tape was presented free-field through Sony speakers at a com-
fortable listening level (centering around 65 dB SPL, as defined by
Davis and Silverman, 1970). A General Radio sound level meter was used
to measure the sound level meter of the tape presentation. The measure-
ments were taken while placing the sound level meter on the table in
front of where the child sat during the test. The loudness attenuator
of the tape recorder was then set at a fixed setting throughout the
testing. All testing was done in a sound-treated suite.

**Concept Teaching**

Each child was tested individually. A set of standard instructions
were used in presenting the task to each child. Each subject was first
familiarized with the concepts "alike" and "not alike". Four pictures
(three smiling girls or boys and one frowning girl or boy) were used to
demonstrate the concepts and were presented with the following direc-
tions:

*See these two girls (or boys)?* (Tester points) *They are both
smiling. They both have brown hair and blue blouses (or shirts).
They look exactly alike.* (Remove pictures and show next two)
*Now look at these two girls (or boys).* (Again pointing) *This
girl is smiling but this girl is frowning (or boy). This girl
is wearing a red blouse and this one is wearing a yellow blouse.
These two girls do not look alike.* (The first set of pictures
was shown again, saying--*) *These girls (or boys) are exactly
alike (and with the second set of pictures--*) *These girls (or
boys) are not alike.*

The child was then asked to point to the pictures that were "alike" and,
next, to the pictures that were "not alike"; each child was verbally
reinforced for a correct response. If the child failed to point to the
appropriate pictures, this same procedure was to be repeated, but all subjects readily pointed to the correct pictures representing "alike" and "not alike"; a few subjects verbalized the words "same" and "different", even though the experimenter did not use these words. The majority of children in this study had completed one year of nursery school and it is possible that they had learned the same-different concept in the nursery classes.

Following the concept teaching trials, each child was told:

First we used our eyes; now we're going to use our ears. I am going to say two words and I want you to listen very carefully. After I say two words, I'll ask you if the two words sounded "alike" (at the same time pointing to the two pictures that were alike) or "not alike" (pointing to the not alike pictures). If the words sounded alike, you say "alike" or put your hand on the alike pictures (demonstrating), but if the two words did not sound alike, you say "not alike" or put your hand on the two pictures that are not alike (demonstrating).

In a pilot procedure, when children were asked to verbalize whether the pairs were "alike" or "not alike", some children continued to point to the pictures representing "alike" and "not alike". Thus, it was decided to provide a choice of responses, either verbal or pointing, in the experimental test procedure.

Task Training

The directions were delivered using the same intonational patterns in presenting the words "alike" and "not alike", to avoid biasing the children's responses. Three preliminary practice pairs of words and nonsense syllables were presented orally until the child's responses indicated that he understood the task. The subjects were presented with three pairs of words and three pairs of nonsense syllables; the
first two pairs contrasted by more than one phoneme, for example, dog-hat and lop-tek, and the last pair contrasted by only one phoneme, for example, bat-bath. They were required to correctly discriminate all six pairs (three pairs of words and three pairs of nonsense syllables) before they were judged to understand the task. If a subject did not correctly discriminate all six pairs, the directions were repeated and six additional practice pairs were used. The subject then had to respond correctly to all six pairs or he was excluded from the study. Two boys were excluded because they did not pass the second familiarization procedure; both failed to attend to the task and either looked around the room or talked.

On all practice items, subjects were verbally reinforced by the examiner for correct responses. The examiner said, "Right, those two words were exactly alike" or "Right, those words were not alike." Similarly, the child was informed of his errors on the practice items. The examiner said, "No, those two words were exactly alike," or "No, those two words were not alike," and then "Now listen carefully to these two words." During the test tapes if the child asked how he was doing or asked to have an item repeated, the examiner commented, "You're doing a good job," or "You're a good listener."

Rest Period

In a pilot procedure with five trial subjects, a three-minute break was provided half way through the test. It was the experimenter's opinion that the children's attention to the task was renewed after the break. Prior to the break there was an increase in nontest behaviors such as, talking and looking around the room. Thus, a three-minute
break between the presentation of the first 48 items (or first 3/5ths) and the last 32 items (or last 2/5ths) was included to provide an opportunity for the child to relax and move around the room. It was not feasible to have a rest period at the half-way point, since the test was developed in fifths; it was arbitrarily decided to break after three fifths were concluded. The instructions explaining how the child was to respond to the test pairs were repeated before each of the two sections on the tape. The examiner faced the child during the familiarization procedures and sat beside the child, out of his line of vision, during the testing to avoid providing differential cues.

Motivation

To maintain motivation and interest throughout the testing, every response made by the child was verbally and tangibly reinforced. During the familiarization procedures and during testing, a glass jar mounted on a wood platform was placed in the child's view but out of his reach. The child was informed that after his every response the experimenter would drop a plastic chip in the jar. It was explained that after all of a certain color of chips were in the jar, he would receive a plastic token; he would receive a second token when all of another color of chips were in the jar. Tokens were exchanged for a reward at the conclusion of the test. The distribution of the colored chips and tokens corresponded with the two parts of the test and the break. During the test, the experimenter's response to each of the child's answers was to smile, say "OK", drop a plastic chip in the jar and then mark the answer sheet.
Scoring

Each response to the test items was scored correct or incorrect by the examiner as it occurred. A "+" was recorded for a correct response and a "-" for an incorrect response. For each test of each child, the total number of errors were computed. This total score was divided into:

(1) the number of errors on "different" or minimal pair items, which was further divided into:

(a) the number of errors in the prevocalic position for both meaningful and nonmeaningful test stimuli, and
(b) the number of errors in the postvocalic position for both meaningful and nonmeaningful test stimuli.

(2) the number of errors on "same" or identical pair items for meaningful and nonmeaningful test stimuli.

After the results of each child's test were tabulated, the test scores of one girl and one boy were eliminated from the study, since neither child made any discriminatory responses on the test. One child indicated all "alike" responses and the other all "not alike" responses to the test items, even though both did pass the task teaching trials. Thus, the study included the responses of 11 girls and 11 boys on the experimental speech sound discrimination test. It was theorized that the following factors may have had an effect on the responses of the two children excluded from the study: (1) during the actual test, the children no longer received a verbal indication of the correctness of their responses (although they were reinforced for any response), and (2) the reinforcement provided may not have been appropriate or effective for these two children.
Statistical Design

All subjects were given a combination word and nonsense syllable discrimination test. The effects of the experimental variables (meaning, position and sex) were evaluated by means of an analysis of variance technique. According to Lindquist (1953, pp. 292-296), this is a Type VI treatment by treatment by subjects by replications statistical design in which each subject takes one level of C (sex) in combination with all possible combinations of A (meaning) and B (position).
CHAPTER III

RESULTS AND DISCUSSION

Introduction

The primary purpose of this study was to investigate the possible effects of the meaningfulness of test stimuli on speech sound discrimination scores of preschool age children. Many different test stimuli have been used to investigate speech sound discrimination, such as words and/or pictures and nonsense syllables. Little auditory discrimination testing has been done with preschool aged children, primarily because most tests are considered too difficult for use with young children. It is hoped that the data obtained in this study will be a helpful addition to the available information on preschool children's auditory discrimination responses.

Fatigue Effect

As described previously, the experimental speech sound discrimination test was sectioned into five equal parts in order to provide for a check on possible fatigue effects. A short break was inserted between the third and fourth sections (fifths) of the test. The total number of errors for each fifth of the test for boys and girls combined were computed and are presented in Appendix B. These total scores per fifth were evaluated for a possible fatigue factor by means of an analysis of variance technique. A Lindquist (1953, pp. 156-158) treatment by
subjects design was used in which the treatments (fifths of a speech sound discrimination test) were all administered in succession to the same subjects. The summary of the analysis of variance of these data is provided in Table 2. The results of this analysis indicated there were no significant differences at the .05 level of confidence between mean scores obtained from each fifth of the speech sound discrimination test. Thus, it was concluded that no fatigue factor was present and the total test scores were retained for analysis.

The "same" pair items were arranged in a regular pattern (every fourth item was a "same" pair). There was no noticeable response pattern in the resulting scores indicating that the subjects had discovered the regular pattern of the "same" pairs. The results of the analysis of variance of the errors per fifth of the test (Table 2) indicated that there was no difference in the number of errors between each fifth of the test and thus supports the contention that there was no evidence that the subjects discovered the regular distribution pattern of the test items (which would have been indicated by higher scores on later fifths than on earlier).

Although no fatigue factor was demonstrated, it was the experimenter's opinion that any future use of this test should continue to include a break providing the subjects the opportunity to move around the room or talk. The same types of nontest behaviors (talking and looking around the room) noted in the pilot procedure, occurred in the experiment prior to the break; there was an obvious reduction in these behaviors after the break. The task of sitting still and listening for 20 to 30 minutes produced an obviously difficult attention factor for
Table 2. Summary of the Analysis of Variance of Errors Per Fifth on the Speech Sound Discrimination Test

<table>
<thead>
<tr>
<th>Source</th>
<th>Degrees of Freedom</th>
<th>Sums of Squares</th>
<th>Mean Squares</th>
<th>F Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatments (Errors/fifth)</td>
<td>4</td>
<td>9,689</td>
<td>2.422</td>
<td>.908</td>
</tr>
<tr>
<td>Subjects</td>
<td>21</td>
<td>289.272</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatments x Subjects</td>
<td>84</td>
<td>223.911</td>
<td>2.666</td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>109</td>
<td>522.872</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

F .01, df 4 and 70 = 3.60
F .05, df 4 and 70 = 2.50
a young child. From observation during the testing, it appeared that
the break together with the motivation and reward techniques, tended to
reduce the occurrence of boredom and fatigue and the analysis summarized
in Table 2 would seem to support this.

Experimental Variables

The data for the evaluation of the effects of the experimental
variables consisted only of the error scores on the minimal paired
items in the test; the error scores on those items consisting of iden-
tical pairs were not included in this analysis. The experimental
speech sound discrimination test consisted of 20 "same" or identical
paired items and 60 "different" or minimal paired items, differing by
one consonant phoneme in either prevocalic or postvocalic positions.
The minimal pair error scores were selected for analysis because:
(1) the effect of the word or syllable position of the contrasted
phonemes could be evaluated only with the minimal pair error scores,
and (2) Wepman's results from administering the Wepman Test of Auditory
Discrimination indicated that errors on "different" or minimal pairs
were significant in determining an auditory discrimination score,
whereas error on "same" pairs were not. The effects of the experi-
mental factors (meaning, position and sex) were evaluated by means of
an analysis of variance technique, a Lindquist Type VI (1953, pp. 292-
297). The analysis involved a consideration of three factors: (1) the
kind of test stimuli: meaningful (word) or nonmeaningful (nonsense
syllable), (2) the position of the contrasted phonemes in the minimal
pairs: prevocalic or postvocalic, and (3) the sex of the subjects: male
or female.
Table 3. Summary of the Analysis of Variance of Scores on an Experimental Speech Sound Discrimination Test

<table>
<thead>
<tr>
<th>Source</th>
<th>Degrees of Freedom</th>
<th>Sums of Squares</th>
<th>Mean Squares</th>
<th>F Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>BETWEEN</td>
<td>21</td>
<td>366.773</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex (S)</td>
<td>1</td>
<td>2.91</td>
<td>2.91</td>
<td>.159</td>
</tr>
<tr>
<td>Error between</td>
<td>20</td>
<td>363.863</td>
<td>18.193</td>
<td></td>
</tr>
<tr>
<td>WITHIN</td>
<td>66</td>
<td>177.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meaning (M)</td>
<td>1</td>
<td>26.182</td>
<td>26.182</td>
<td>12.642</td>
</tr>
<tr>
<td>Position (P)</td>
<td>1</td>
<td>.728</td>
<td>.728</td>
<td>.2073</td>
</tr>
<tr>
<td>M x P</td>
<td>1</td>
<td>2.225</td>
<td>2.225</td>
<td>1.283</td>
</tr>
<tr>
<td>M x S</td>
<td>1</td>
<td>.408</td>
<td>.408</td>
<td>.197</td>
</tr>
<tr>
<td>P x S</td>
<td>1</td>
<td>.407</td>
<td>.407</td>
<td>.116</td>
</tr>
<tr>
<td>M x P x S</td>
<td>1</td>
<td>.733</td>
<td>.733</td>
<td>.423</td>
</tr>
<tr>
<td>Error within</td>
<td>60</td>
<td>146.317</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error(_1)</td>
<td>20</td>
<td>41.41</td>
<td>2.071</td>
<td></td>
</tr>
<tr>
<td>Error(_2)</td>
<td>20</td>
<td>70.227</td>
<td>3.511</td>
<td></td>
</tr>
<tr>
<td>Error(_3)</td>
<td>20</td>
<td>34.68</td>
<td>1.733</td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>87</td>
<td>543.773</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

F\(_{.01}\), df 1 and 20 = 8.10
F\(_{.05}\), df 1 and 20 = 4.35

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
The summary of the analysis of variance of these data is in Table 3. Table 4 displays the mean number of errors for each of the three experimental variables (meaning, position and sex) on the experimental test of discrimination. All of the values for the various statistical analyses were obtained from the raw data presented in Appendix A. The results of the analysis of variance indicate that there are statistically significant differences at the .01 level of confidence between the means of the scores on the meaningful and nonmeaningful test stimuli. While meaningfulness had the predicted effect on discrimination scores (fewer errors on meaningful material), there was no statistically significant differences (at the .05 level of confidence) associated with any of the other main effects (position or sex), nor were any of the interactions significant.

Table 4. Mean Number of Errors for Meaning, Position and Sex on the Experimental Test of Discrimination

<table>
<thead>
<tr>
<th>Meaning</th>
<th>Nonmeaning</th>
<th>Prevocalic</th>
<th>Postvocalic</th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.614</td>
<td>3.705</td>
<td>3.068</td>
<td>3.25</td>
<td>3.341</td>
<td>2.978</td>
</tr>
</tbody>
</table>

Meaningfulness of Test Stimuli

It would appear that it is more difficult for preschool aged children with normal hearing and articulation skills to discriminate between phonemes when meaning is removed from the test items. Since children scored higher on the meaningful or word items, this would suggest that a child's familiarity with the language or his recognition of the word test items may be a critical feature in his discriminative
judgments. Vocabulary or the recognition of words appears to be a factor in this and other investigations. Elenbogen and Thompson, in a study previously cited, found that it was more difficult for kindergartners of middle class backgrounds to discriminate between phonemes when meaning was removed from the test items, but this did not appear to be true for kindergartners of lower class backgrounds. They concluded that a vocabulary factor was measured along with speech sound discrimination in the word items on the measure used (Wepman Test of Auditory Discrimination). Thus, meaningfulness of the stimuli may add an additional contextual cue to aid the child in discriminating between test items. Nonsense syllables, being relatively devoid of meaning, do not contain the cues to discrimination which are provided in meaning in the recognition of words.

The strength of the relationship between scores on the word and nonsense syllable test items was evaluated by the Pearson Product Moment Correlation Coefficient. The correlation coefficients are reported in Table 5. The correlations indicated a moderate relationship between scores on word and nonsense syllable items on the experimental speech sound discrimination test. It would appear that there is a moderate degree of similarity in the discrimination tasks using words and nonsense syllables; however, the coefficient of determination indicates that the scores on the two types of test stimuli are more independent than dependent. Thus, a moderately related and yet importantly independent skill or skills were measured when the two test stimuli (words and nonsense syllables) were used. Factors, such as the age of the subjects and their development of discrimination skill and the
Table 5. Correlations Between Scores for the Variables of Meaning and Position

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficient of Correlation</th>
<th>Coefficient of Determination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meaning and Nonmeaning Scores</td>
<td>.512</td>
<td>.262</td>
</tr>
<tr>
<td>Prevocalic and Postvocalic Positions on Meaningful Items</td>
<td>.815</td>
<td>.664</td>
</tr>
<tr>
<td>Prevocalic and Postvocalic Positions on Nonmeaningful Items</td>
<td>.586</td>
<td>.344</td>
</tr>
</tbody>
</table>
differences in the familiarity of the test items are some possible sources of variability in discrimination test scores beyond meaningfulness of test items.

In the literature it has been suggested that a child learns to discriminate between sounds as he learns the words of his language. The results of this study would indicate that a child learns a specific rather than a general discrimination ability, i.e., he learns to discriminate between sounds within specific words and the discrimination ability does not generalize to all speech stimuli (as demonstrated by more errors on nonmeaningful items).

The development of discrimination skill is taking place within the context of language learning. This is not to say that the more words a child learns, the more he is able to discriminate; Templin (1957) found very low correlations between discrimination and vocabulary scores. Rather, discrimination skill learning takes place concurrently with language acquisition and is practiced in the course of using and understanding oral language. It appears that, as test items, words provide more "usual" stimuli for young children for the task of discriminating between phonemes. Nonsense syllables, as test stimuli, would seem to require an application or transfer of discrimination skill, a skill which the child aged 4 to 5 years is still developing (according to Templin, 1957). Possibly with older children, scores on nonsense syllable items would provide a way to differentiate between levels of discrimination skills among subjects who might do equally well on a word test.

Templin (1957) developed a speech sound discrimination test for preschool children (ages 3 - 5) which involved matching an auditorily
presented stimulus with a picture; a nonsense syllable discrimination test was designed to be used with school aged children or older. Templin stated that the difficulty of making judgments between abstract nonsense syllables restricted the use of the nonsense syllable test to elementary grades and above. Since more errors occurred on the nonsense syllable items of this experimental discrimination test, it could be argued that the nonsense syllable items did increase the difficulty and/or abstractness of the test. However, the experimenter would argue Templin's statement that nonsense syllable items should be restricted to use with school aged children. Considering the total number of errors occurring on the nonsense syllable items of the experimental discrimination test (207 errors out of a possible 880), the subjects failed only 23.5 percent of the nonsense syllable items, whereas they failed 18.2 percent of the word items (161 errors out of a possible 880). Thus, in this sample of subjects, the use of nonsense syllable items in discrimination testing with preschool aged children increased somewhat the number of errors when compared with errors on word items, but the results of this experiment demonstrated that the task of making such judgments is well within the capability of the child aged 4½ to 5½ years. Again, the somewhat higher ceiling on nonsense syllable test scores makes it possible to differentiate among older children, all of whom may do equally well on a word test. For future research, then, it would be of interest to compare groups of older children on the experimental speech sound discrimination to note the possible effects of the meaningfulness of the test stimuli.

The experimenter's intent in this study was to use "average" children as subjects in an exploratory study to obtain data on speech
sound discrimination ability of preschool children. Hearing acuity and articulation skills were controlled to avoid biasing the sample. Since previous research has indicated that children of lower socioeconomic status have more articulation errors than children of upper socioeconomic status, it was assumed that screening for articulation skill would indirectly reduce the effect of socioeconomic status; thus direct control or identification of socioeconomic status was not utilized in this study. Although children with "average" hearing and articulation skills made fewer errors on the meaningful items in this study, an advantage to the use of nonsense syllables as test items is the reduction of a vocabulary or language learning factor present in word items. This would seem important when the goal would be the early identification of children with speech sound discrimination difficulties without contaminating the testing with vocabulary factors or exposure to test items. The use of nonsense syllables may provide appropriate discrimination test items for testing and comparing children with different levels of language skills. Thus, for further research, it would also be of interest to compare two groups of children, either from different socioeconomic backgrounds or at different measured levels of language skill, on this experimental discrimination test.

Position of Sound Contrasts

The predicted influence of position (fewer errors on sound contrasts in the prevocalic position) was not significant. The results indicated that neither the prevocalic or postvocalic position was more difficult to discriminate between sound contrasts. This finding does not necessarily support the contention that a consonant sound in either
the prevocalic or postvocalic position involves identical discriminatory events. The production of most individual consonants is modified when in contact with other sounds (Malmberg, 1963). When discriminating a phoneme in both prevocalic and postvocalic positions, that phoneme itself has been modified by the sound context; thus, it cannot be said that the discrimination of one sound contrast is an identical task in different sound positions. This study indicates that whatever effect the position of the sound contrast may have had, it did not make a significant difference in the resulting discrimination scores.

Templin (1943), in a study of speech sound discrimination ability of elementary school pupils, grades two through six, found that sound contrasts in the medial or final (postvocalic) positions were more difficult to discriminate than initial (prevocalic) position sound contrasts. The results of the present study (no position effect), contrast with Templin's 1943 findings (initial position contrasts are easier to discriminate). There are differences in the procedures of both studies which may account for this difference in results. Templin's test used all nonsense syllable test items, whereas the experimental discrimination test used an equal number of words and nonsense syllable items. Templin administered all 100 initial (prevocalic) position items followed by 100 medial or final (postvocalic) position items. It is possible that the order and/or length of the test may have had an effect on the resulting scores in Templin's study and that fatigue effects interacted with sound-position effects. In the present study, the position of the sound contrasts was randomized within each fifth of the experimental discrimination test and the results were analyzed for a
possible fatigue factor and none was evident. The ages of the subjects in the two studies differed: 8 to 12 years for Templin's study and $4^{1/2}$ to $5^{1/2}$ for the present study. Thus, the following factors may have influenced the difference in sound contrast position results: test items, subject age, and length of the test. Kamil and Rudegeair (1972) found that after administering similar forms of a discrimination test on successive days, there were no significant differences in scores with regard to the position of the sound contrasts. As mentioned previously, the results of the effect of the sound position during the first administration of the test were not reported, thus making it difficult to compare their results with Templin's or the present study.

One could speculate that differences in test items between Templin's and the present study had an effect, such that discriminating sounds in nonsense syllables is a more difficult task than discriminating sounds in words. Again, the meaningfulness of the word items in the experimental test may have aided in the discrimination task. Possibly as the abstractness of the items increases (more nonsense syllable items), there is a greater difference in the effect of position; the meaning by position interaction (Table 3) yielded the largest F ratio among the interactions, though it did not reach conventionally acceptable levels of significance.

A correlation coefficient was computed to measure the strength of the relationship between the prevocalic and postvocalic positions in the meaningful and nonmeaningful test items. These correlations are reported in Table 5. The correlation (.815) between prevocalic and postvocalic sound positions for meaningful items indicated a moderately
strong relationship between scores for items in these two sound positions. The correlation (.586) between prevocalic and postvocalic sound positions for nonmeaningful items indicated only a moderate relationship between scores for items in these positions. The correlations demonstrated a lower relationship between scores of prevocalic and postvocalic positions when using nonmeaningful material than when using meaningful material. Again, this supports the contention for a possible meaning by position interaction, even though such an interaction was not significant in this study.

For older children, Templin's discrimination results correspond with her articulation results; that is, more articulation errors occurred in the final (postvocalic) sound position than in the initial (prevocalic) position and the final (postvocalic) position contrasts were more difficult to discriminate than initial (prevocalic) position contrasts. It is possible that with the child aged 4 to 5 years there is not a clear, direct relationship between producing a particular sound in a specific word position and auditorily discriminating that same sound from another sound in the same word position, as indicated by Templin's articulation and discrimination findings for older children.

**Sex of the Subjects**

As predicted, the differences in the mean scores between the sexes were not statistically significant. However, the boys did score slightly higher than the girls, the opposite of which had been predicted. Templin (1957) administered a word-picture discrimination test to children aged three through five years. Although Templin's test and
the experimental test were different in structure, similar discrimination results were obtained. Templin found that, while girls scored slightly higher than boys from age four on, the sex differences were not statistically different. Templin (1957) administered a nonsense syllable test to children aged six through eight years. For this entire age range, the mean scores of the sexes were significantly different, with girls scoring higher than boys. It is possible that the development of discrimination skill follows a similar pattern as the development of articulation skills. In an investigation of articulation skill, Templin (1957) found that, although girls produced more sounds correctly than did boys, this difference was not significant until age seven. Thus, significant sex differences have been reported for both discrimination and articulation skills after six years of age. The subjects in the present study ranged in age from $4\frac{1}{2}$ to $5\frac{1}{2}$ years. It appears that at this age ($4\frac{1}{2}$ to $5\frac{1}{2}$ years), boys and girls are developing discrimination at about the same rate. From the research cited above, it seems possible that as a child reaches the ages of six through eight years, one could predict the occurrence of a significant difference in the discrimination scores, with girls scoring higher.

### Analysis of "Same" Pairs

Twenty-five percent of the experimental test items were "same" or identical pairs. A separate analysis of variance was used to evaluate the total number of errors on the "same" or identical pairs. A Type I (Lindquist, 1953), mixed two-factor analysis of variance design in which each of the A treatments (meaningful or nonmeaningful test stimuli) was administered with either of the B treatments (sex of the
subject), was used to evaluate the "same" pair errors. The summary of this analysis of variance is provided in Table 6, and it indicates that there are no statistically significant differences between the means of the test stimuli with regard to errors on "same" pairs. Thus, performance on "different" or minimal pair items yielded significant differences with regard to the meaningfulness of the test stimuli while performance on "same" or identical pair items yielded no significant difference with regard to meaningfulness. This difference in results for "different" and "same" pair items is consistent with Wepman's findings that scores on "different" pairs are significant in determining an auditory discrimination score while scores on "same" pairs are not. For this study, scores on "different" items are significant in measuring the effect of meaningfulness whereas the "same" items are not significant.

The greater skill in detecting differences between phonemes in meaningful material than in nonmeaningful might imply a greater willingness to hear different phonemes as the same when they occur in a nonmeaningful context. Such a "same" response bias for nonsense syllables was not borne out by a comparison of the percentages of errors on different pairs versus that on same pairs for meaningful and nonmeaningful materials (Table 7). The children made essentially equal proportions of error responses for both same and different pairs of nonsense materials and, thus, a "same" response bias for nonmeaningful material could not be supported by the data from this study. Apparently these children had more difficulty correctly identifying sameness and differentness of phonemes occurring in nonsense material.
Table 6. Summary of Analysis of Variance of "Same" Pair Errors

<table>
<thead>
<tr>
<th>Source</th>
<th>Degrees of Freedom</th>
<th>Sums of Squares</th>
<th>Mean Squares</th>
<th>F Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>BETWEEN</td>
<td>21</td>
<td>134.818</td>
<td>6.119</td>
<td></td>
</tr>
<tr>
<td>Sex (S)</td>
<td>1</td>
<td>1.363</td>
<td>1.363</td>
<td>.204</td>
</tr>
<tr>
<td>Error between</td>
<td>20</td>
<td>133.455</td>
<td>6.672</td>
<td></td>
</tr>
<tr>
<td>WITHIN</td>
<td>22</td>
<td>67.0</td>
<td>3.045</td>
<td></td>
</tr>
<tr>
<td>Meaning (M)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M x S</td>
<td>1</td>
<td>.09</td>
<td>.09</td>
<td>.0269</td>
</tr>
<tr>
<td>Error within</td>
<td>20</td>
<td>66.91</td>
<td>3.345</td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>43</td>
<td>201.818</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$F_{.01}$, df 1 and 20 = 8.10

$F_{.05}$, df 1 and 20 = 4.35

Table 7. Percentages for Total Number of Errors on Meaningful and Nonmeaningful Test Stimuli

<table>
<thead>
<tr>
<th>Different Pair Errors</th>
<th>Same Pair Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meaningful</td>
<td>Meaningful</td>
</tr>
<tr>
<td>Nonmeaningful</td>
<td>Nonmeaningful</td>
</tr>
<tr>
<td>18.2</td>
<td>20.9</td>
</tr>
<tr>
<td>23.5</td>
<td>20.0</td>
</tr>
</tbody>
</table>
Recommendations

The intent of this study was to obtain data on the possible effects of the meaningfulness of test stimuli on speech sound discrimination scores of preschool children. Children aged $4\frac{1}{2}$ to $5\frac{1}{2}$ years, with normal hearing and articulation skills were chosen for subjects. Since only 22 subjects were used in the final analysis, it is difficult to generalize to other populations. Further research should include larger samples. During the above discussion, a number of possibilities for future research and investigations for normative data were noted. These include the following:

1. The experimental speech sound discrimination test could be administered to different age groups to note the effects of meaningfulness of test items, position and sex as age increases.

2. It has been postulated that the nonsense syllable items may serve to differentiate among subjects, all of whom may do equally well or "ceiling out" on a word discrimination test. The effect of meaningfulness could be investigated with older children by administering the experimental speech sound discrimination test.

3. The experimental discrimination test could be administered to children with different socioeconomic backgrounds or different measured language skills. Again, the effects of meaningfulness, position and sex could be investigated.

4. The experimental discrimination test could be administered to a large group of children, and rather than control age
or group children according to age levels, children would be assigned to groups according to the score obtained on the experimental test. Then the meaningfulness, position and sex effects could be investigated for groups of "good" and "poor" discriminators.

5. Although it was noted earlier that all test items in the experimental test had some relative degree of meaningfulness, the items were classified as either meaningful or nonmeaningful. The experimenter would agree that meaningfulness exists in degrees. In terms of future research, the meaningfulness of words for young children should be investigated. The meaningfulness of words could be measured by a satiation technique or a scaling technique.
CHAPTER IV

SUMMARY AND CONCLUSIONS

An investigation was conducted to determine whether the meaningfulness of the test stimuli used in an experimental speech sound discrimination test (composed of word and nonsense syllable test items) would produce statistically significant differences in the resulting scores of a group of preschool aged girls and boys.

Twenty-four children, 12 boys and 12 girls, ranging in age from four years-six months to five years-six months, with normal hearing and articulation skills, were chosen as subjects for this study. Each child was given the experimental speech sound discrimination test individually in a sound-treated environment.

The results obtained were evaluated by means of an analysis of variance technique. The analysis involved a consideration of the following factors:

1. Meaningfulness of test stimuli (M)
2. Position of the sound contrasts (P)
3. Sex of the subjects (S)

The results of the analysis of variance indicated that the predicted differences between the means on the meaningful and nonmeaningful test items was statistically significant at the .01 level of confidence; there were no statistically significant differences at the .05 level.
of confidence associated with any of the other main effects (position or sex), nor were any of the interactions significant.

Recommendations for further studies of a similar nature were discussed.
BIBLIOGRAPHY
BIBLIOGRAPHY


Noble, C. E. "Measurements of Association Value (a), Rated Associations (a) and Scaled Meaningfulness (m) for the 2100 CVC Combinations of the English Alphabet," Psychological Reports, 8, 1961, 487-521.


### APPENDIX A

**SUMMARY OF RAW DATA: ERROR SCORES**

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Word Initial</th>
<th>Word Final</th>
<th>Word &quot;Same&quot;</th>
<th>Nonsense Initial</th>
<th>Nonsense Final</th>
<th>Nonsense &quot;Same&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Females:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>10</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>5</td>
<td>8</td>
<td>1</td>
<td>4</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td><strong>Males:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>6</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>7</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>6</td>
<td>0</td>
<td>5</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>
APPENDIX B

TOTAL NUMBER OF ERRORS PER FIFTH OF THE EXPERIMENTAL SPEECH DISCRIMINATION TEST FOR GIRLS AND BOYS COMBINED

<table>
<thead>
<tr>
<th>Fifths</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total errors</td>
<td>76</td>
<td>83</td>
<td>69</td>
<td>64</td>
<td>76</td>
</tr>
</tbody>
</table>
APPENDIX C

EXPERIMENTAL SPEECH SOUND DISCRIMINATION TEST

1. sick-sit
2. pæf-bæf
3. shop-chop
4. leaf-leaf
5. tæθ-tæs
6. wait-rate
7. zæf-zæf
8. fæθ-fræθ
9. wish-witch
10. wop-wob
11. cook-took
12. sing-sing
13. veθ-veθ
14. raids-rage
15. θæp-sæp
16. lit-lit
17. fought-thought
18. mærd-nærd
19. beg-bed
20. hide-hide
21. wk-stk
22. mash-mass
23. lob-lov
24. dæs-dæs
25. gate-date
26. sæk-sæk
27. deaf-death
28. bus-bus
29. tots-to
30. shoot-suit
31. bok-vok
32. harb-harb
33. path-bath
34. kæθ-kæθ
35. thumb-some
36. full-full
37. sæk-sæθ
38. safe-save
39. dæk-ðæk
40. dæθ-tæθ

41. rope-robe
42. jæθ-jæθ
43. vase-face
44. jæm-jæm
45. tæθ-tæθ
46. pass-path
47. woθ-woθ
48. sul-sul
49. might-night
50. gek-dek
51. coats-coach
52. walk-walk
53. fæp-θæp
54. boat-vote
55. næs-næs
56. wæθ-wæθ
57. wet-let
58. gæθ-gæθ
59. them-then
60. dish-dish
61. gæθ-gæθ
62. robe-rove
63. jæk-sæk
64. kæθ-kæθ
65. sink-zinc
66. weg-reg
67. bed-bet
68. pig-pig
69. jæθ-jæθ
70. load-loathe
71. væk-væθ
72. jæk-jæk
73. time-dime
74. ledz-læθ
75. rise-rice
76. yet-yet
77. læk-læθ
78. dare-there
79. kvæp-tæp
80. tæθ-tæθ