Eastern Illinois University
The Keep

Masters Theses

Student Theses & Publications

1973

# A Method of Improving the Predictive Validity of the Predictive Screening Test of Articulation

Bradley R. Chase
*Eastern Illinois University*

This research is a product of the graduate program in Communication Disorders and Sciences at Eastern Illinois University. Find out more about the program.

Recommended Citation

Chase, Bradley R., "A Method of Improving the Predictive Validity of the Predictive Screening Test of Articulation" (1973). *Masters Theses*. 3737.
https://thekeep.eiu.edu/theses/3737

This is brought to you for free and open access by the Student Theses & Publications at The Keep. It has been accepted for inclusion in Masters Theses by an authorized administrator of The Keep. For more information, please contact tabruns@eiu.edu.

A METHOD OF IMPROVING THE PREDICTIVE VALIDITY OF

THE PREDICTIVE SCREENING TEST OF ARTICULATION

(TITLE)

BY

Bradley R. Chase

**THESIS**

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

Master of Science

IN THE GRADUATE SCHOOL, EASTERN ILLINOIS UNIVERSITY

CHARLESTON, ILLINOIS

1973

YEAR

I HEREBY RECOMMEND THIS THESIS BE ACCEPTED AS FULFILLING

THIS PART OF THE GRADUATE DEGREE CITED ABOVE

2/22/73

DATE

ADVISER

Feb 22, 1973

DATE

PAPER CERTIFICATE #2

TO: Graduate Degree Candidates who have written formal theses.

SUBJECT: Permission to reproduce theses.

The University Library is receiving a number of requests from other institutions asking permission to reproduce dissertations for inclusion in their library holdings. Although no copyright laws are involved, we feel that professional courtesy demands that permission be obtained from the author before we allow theses to be copied.

Please sign one of the following statements.

Booth Library of Eastern Illinois University has my permission to lend my thesis to a reputable college or university for the purpose of copying it for inclusion in that institution's library or research holdings.

2-22-73
Date

I respectfully request Booth Library of Eastern Illinois University not allow my thesis be reproduced because ______________________________

______________________________________________________________

______________________________________________________________

______________________ ______________________
Date Author

ACKNOWLEDGMENTS

The writer wishes to acknowledge the interest and assistance of the following persons:

Acknowledgment is given to Dr. Lynn Miner for serving as graduate and thesis advisor and for supervision of the formulation of the research design and the statistical analysis of data.

Acknowledgment is given to Dr. Wayne Thurman, Dr. Jerry Griffith, and Mrs. Mary Beth Armstrong for their supervision and advice as members of the graduate committee.

Acknowledgment is given to Dr. Robert Erickson and Dr. Charles Van Riper for allowing the writer to have access to their research data.

A very special thanks is given to my lovely wife, Mrs. June (Imberman) Chase, for her assistance in typing the thesis. This project is dedicated to her.

TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## CHAPTER I

## INTRODUCTION

Public school speech correctionists are presented with a serious problem when selecting caseloads for the school year. They must continuously decide whether or not a child with an articulatory defect in kindergarten or the first grade will correct his misarticulations on his own through maturation, or will require special speech training to correct his errors. The problem, then, is that the speech clinician may give his time to a child who really does not need it, and children who actually need his special attention are neglected.

Speech clinicians in the clinical setting are faced with a similar problem. When a child is given a speech evaluation, the clinician often has difficulty deciding whether or not to recommend therapy since no efficient and reliable method of predicting an individual child's articulatory improvement is now available.

The problem is likely to be even more complex for the speech pathologist in private practice. He also has an obligation to treat clients who are in actual need of his services, especially since the speech pathologist is paid directly by the client in most cases. Should the pathologist decide that a child with misarticulations does need therapy when actually the errors will be corrected through maturation, he will have collected fees for, at best, speeding the process of articulatory maturation.

Presently, norms are available for comparison of an individual child's articulatory performance with the performances of his peers. It is likely that the norms collected by Poole (1934) and Templin (1957) are those most commonly used by the speech clinician in evaluation of a child's articulatory skills. Sander (1972), however, has discussed the expected variability of speech sound acquisition among children and pointed out that these normative data represent arbitrarily chosen upper age limits rather than average performances, and stress sound mastery in all word positions as opposed to customary usage.

Furthermore, Sander (1972) summarized age information and presented a table in which sounds were assigned to the earliest ages at which they were correctly produced more than 50 percent of the time in all three positions combined by 51 percent or more children.

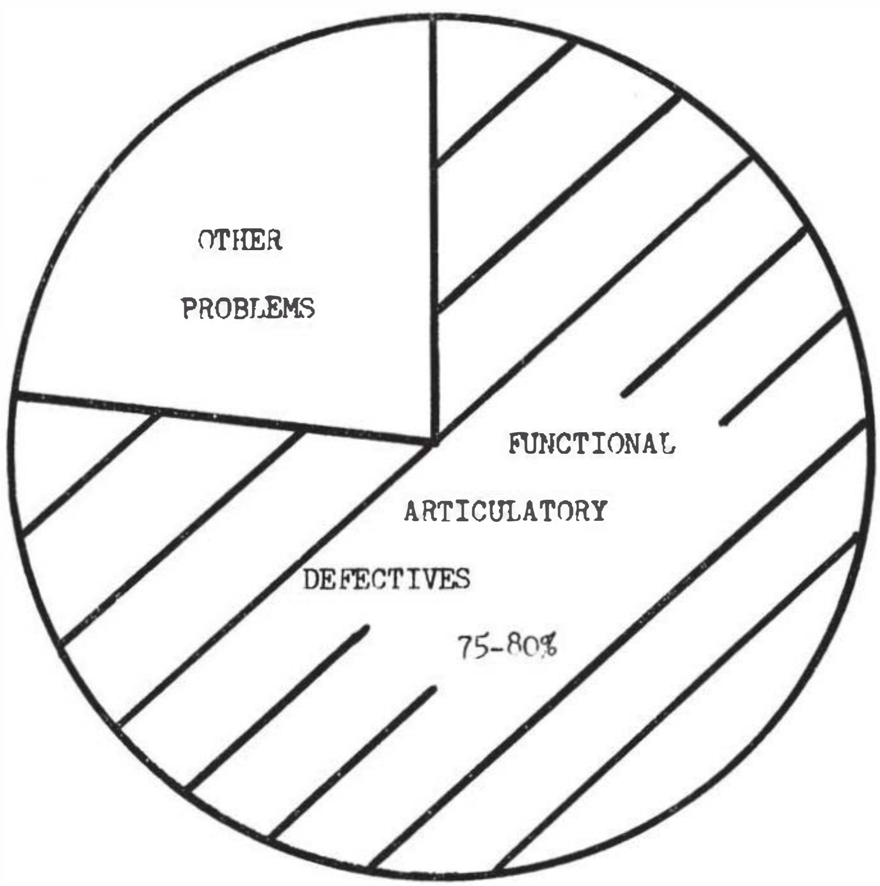
The late Hall-Healey (1963, 1964) norms used a 90 percent criterion of correct production; therefore, the observations set forth by Sander of the Poole and Templin norms are likely to apply to them as well. The basic problem with the age information discussed here is that the clinician may observe articulatory performances of children and compare them with these kinds of tables; however, there are no guidelines for predicting whether or not children who display misarticulations will attain correction at an older age without speech therapy.

Root (1926) reported that of the children with speech defects in the schools of South Dakota, 60 percent had articulatory problems. Loutitt and Halls (1936) found that of the school children whom they surveyed, 79 percent had articulatory problems; and, Reid (1947b) concluded that at least 50 percent of the speech defectives in elementary schools had functional articulatory problems. More recent statistics by Bingham, et al. (1961, p. 38) from a nationwide sample of clinicians, reported that 81 percent of their caseloads was composed of functional articulatory problems; and, Powers (1971, p. 839) has suggested that functional articulatory defectives represent between 75 and 80 percent of all speech defectives in the school population (See Figure 1, p. 4).

From these figures, it appears that a large number of speech defectives possess functional articulatory problems. Arey (1938) listed crowded schedules in the public schools as a limiting factor in speech education programs. Reid (1947b) observed that some children did not attain perfect articulation during their elementary school years without special training; however, the ones who would could not be predicted with present knowledge and methods.

FIGURE 1

GRAPH REPRESENTING THE INCIDENCE OF FUNCTIONAL ARTICULATORY DEFECTIVES IN THE CASELOADS OF PUBLIC SCHOOL SPEECH CORRECTION PROGRAMS



Before and after that observation was made by Reid in 1947, attempts at the construction of prognostic tests of articulation have shown predictive potential, but none have reached high levels of predictive validity.

Indeed, it is necessary that a predictive test be constructed which will aid in the selection of articulatory defective children in the caseloads of the public school speech correctionist. If the incidence of children with articulatory problems within the schools is as high as reported, then a valid predictive test can be of great value.

# CHAPTER II

# REVIEW OF RELATED LITERATURE

## Speech Sound Acquisition and Development

A great amount of information is available concerning the rate of speech sound acquisition among children. Progressive mastery of the correct articulation of phonemes has been demonstrated by Metraux (1950), Poole (1934), Root (1926), Roe and Milisen (1942), Templin (1957), and others. Farquhar (1961) has demonstrated that children with mild and severe articulatory disorders improved their skills without speech therapy over a seven-month period.

Not only is speech sound acquisition a gradual process which occurs with maturation, but Irwin (1947), and Irwin and Curry (1941) have observed that there is an orderly progression of sound acquisition. Vowels appear to develop from front to back, and consonants develop in a similar manner, labials and post-dentals being the earliest.

The child appears to master his articulatory skills until he reaches a fairly definite age. Studies by Loutitt and Halls (1936), Mills and Streit (1942), and Roe and Milisen (1942) showed a marked decrease in the percentages of children with articulatory defects from grades one through three, and a similar decrease in number of errors per child. Even more specifically, Roe and Milisen (1942) found significant differences in mean number of errors between children in grades one and two, two and three, and three and four. Dawson (1929) found that a child's ability to articulate sounds developed after they entered school, and that development was greatest during the first three years of school. The research by Poole (1934) and Templin (1957) demonstrated that children improve their articulation without therapy as they grow older, until they reach the fourth grade. This work is supported by Saylor (1949) and Roe and Milisen (1942) who found that only slight and inconsistent improvement occurred in articulation among students farther along than the third or fourth grades.

From a recent national speech and hearing survey, Williams (1971) has reported that the incidence of misarticulations decreases as grade level increases. The decrease in misarticulations was most profound from grades one to four (See Figures 2 and 3, pp. 8 and 9).

It has been clearly indicated that spontaneous articulatory mastery occurs until the child reaches about the fourth grade. It has been suggested that the most significant improvement occurs during the first through third grades.

FIGURE 2

MEAN GOLDMAN-FRISTOE ERRORS BY GRADE AND SEX

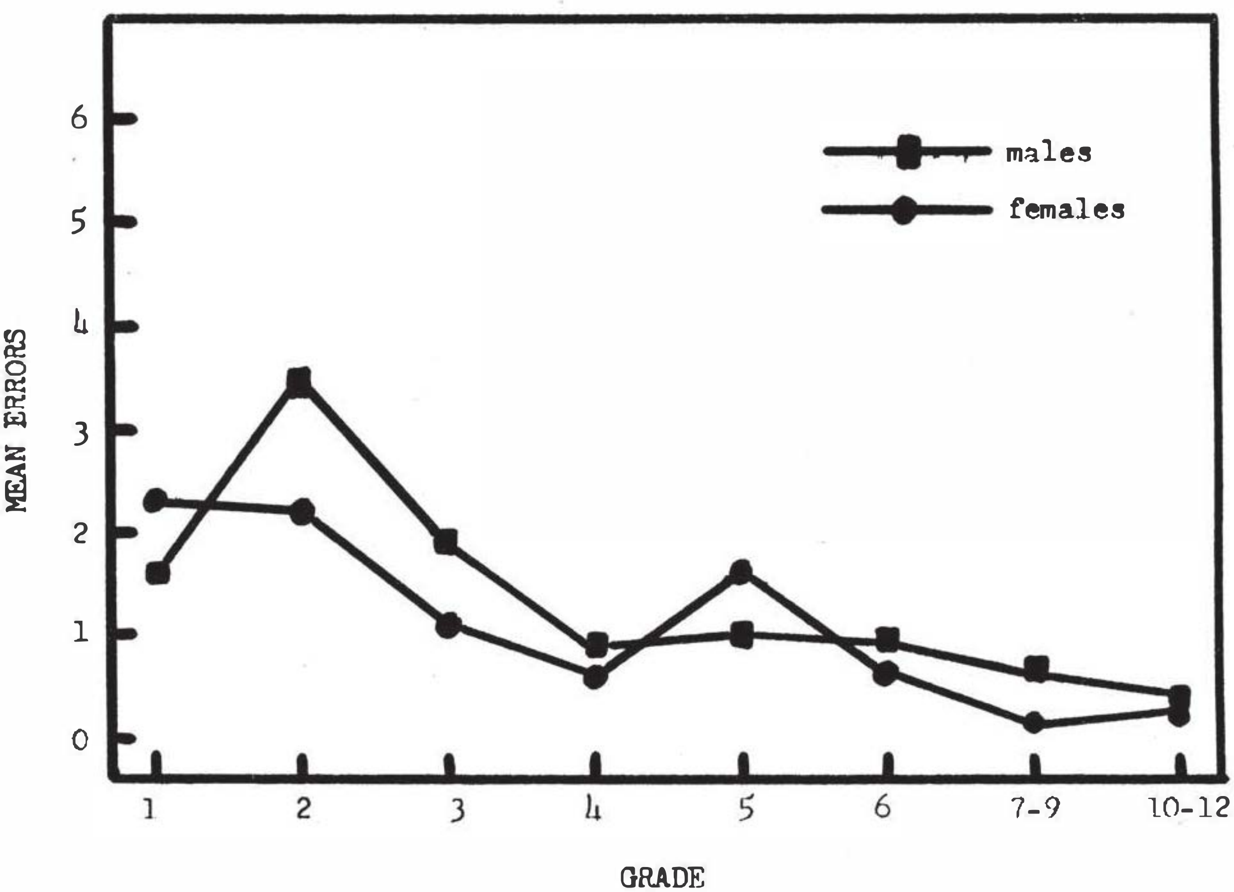
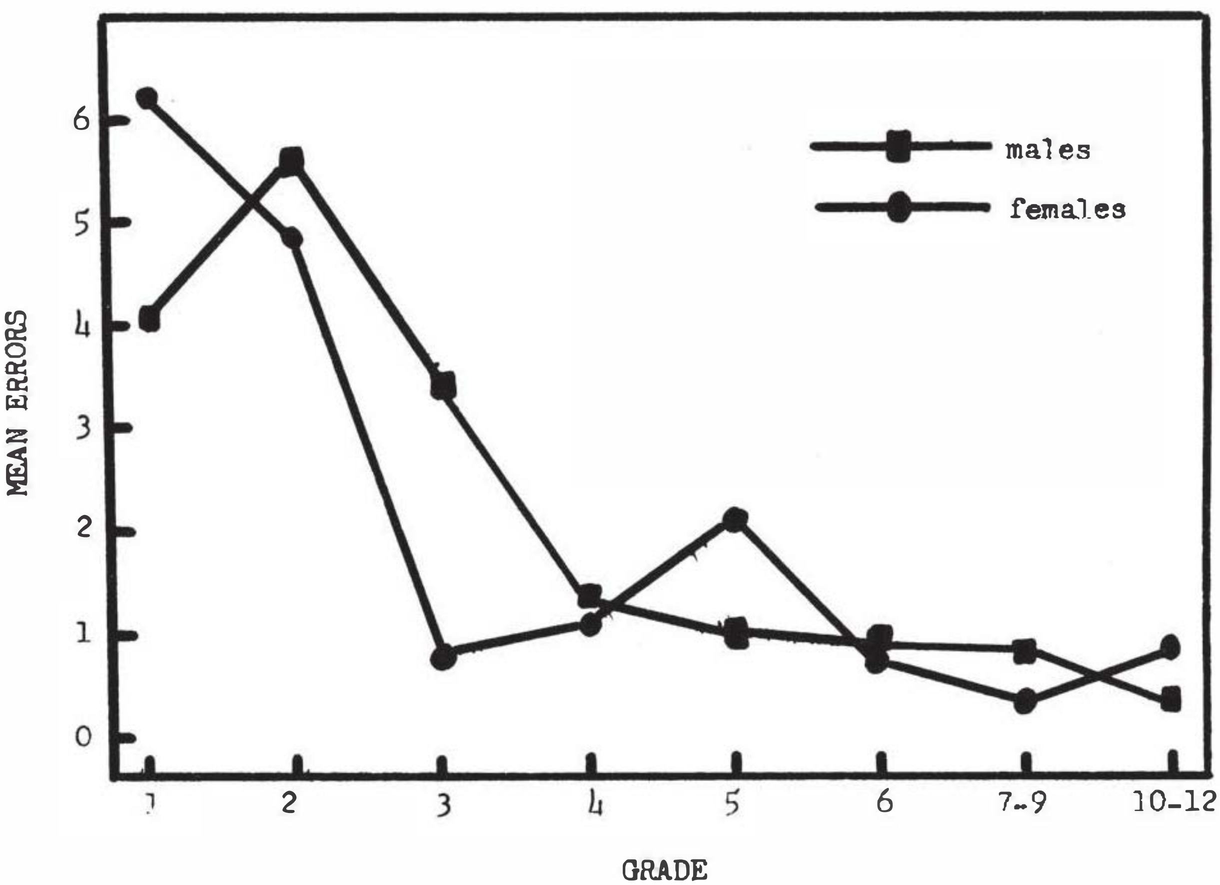


FIGURE 3

MEAN FREE SPEECH ERRORS BY GRADE AND SEX



## Imitation and Spontaneous Articulation Testing

A common method of evaluating a child's articulatory skills is to present the child with pictures or objects designed to elicit the stimulus word featuring the test sound in one or more positions. This method serves to elicit spontaneous utterances from the child. The clinician may then compare these utterances with those of the child's peer group. If it is the clinician's goal to find out whether or not the child is performing comparably to his peers at the present time, this can easily be done. However, if the clinician wishes to know if the child who is developing more slowly than his peers will catch up to them at a later age level, then a spontaneous articulation test does not yield that kind of information.

It has been suggested by Farquhar (1961), Snow and Milisen (1954b), Nichols (1964), and Carter and Buck (1958) that a child's ability to imitate specific sounds and to correct misarticulations after auditory stimulation may be a strong indication that spontaneous improvement will occur.

The children in studies by Carter and Buck (1958) and Snow and Milisen (1954a) showed significant improvement in articulation test scores with imitative testing over spontaneous testing. Similar results were obtained by Smith and Ainsworth (1967) who used combined auditory-visual stimulation in their study of first-grade children with speech defects, and by Siegel, Winitz, and Conkey (1963) who tested kindergarten children not considered to be defective articulators. Farquhar (1961), in her investigation of the articulatory responses of kindergarten children with mild and severe articulatory defects, found that the "mild" group was superior to the "severe" group in imitating sounds in isolation, nonsense-syllables, and in words.

Furthermore, research by Scott and Milisen (1954b) and Smith and Ainsworth (1967) suggests that combined auditory-visual stimulation elicits more correct responses than either visual or auditory stimulation alone. Scott and Milisen (1954b) also found that sounds with visible focal articulation points and sounds with focal articulation points that can be made more visible were most stimulable.

As it has been reported, it appears reasonable to expect that children will improve their articulatory responses after either auditory or combined auditory-visual stimulation. Farquhar (1961) has concluded that imitation of words and nonsense-syllables may be used as prognostic tools. Also, Snow and Milisen (1954b) have concluded that the differences in a child's responses to an oral and a picture articulation test could be used as a valuable factor in predicting his prognosis in correcting his articulation errors. Similar findings by Carter and Buck (1958) suggest that a child's ability to correct his articulation errors instantaneously is indicative of the degree of speech maturation. Consequently, this is a strong suggestion for the development of a prognostic articulation test which lends itself to predictive scoring.

Templin (1947a), however, tested the effects of spontaneous versus imitation testing on 100 children ranging in age from two years, one month to five years, nine months. She concluded that in measuring total articulation of preschool children, similar results were obtained if the sounds were measured in words spontaneously elicited, or repeated after the examiner. The mean age of the subjects in Templin's study was three years, nine months with a standard deviation of 8.9 months. It is possible that Templin's results disagree with the findings of the other investigators mentioned since the subjects used in the other studies were of kindergarten age or older. The relationship between chronological age and the effects of imitation testing seems to be an appropriate subject for further research.

From the majority of these findings, it seems reasonable that any valid predictive articulation test will necessarily involve imitation testing, since it enables the clinician to evaluate a child's ability to produce sounds after auditory or auditory-visual stimulation. If it is correct to believe that children learn speech and acquire speech sounds by hearing auditory signals and imitating them, then imitative testing is understandably a valuable prognostic tool.

## Auditory Discrimination and Articulation

Although only conflicting information is available concerning the relationship between speech sound discrimination and articulation, speech sound discrimination may possibly be an important consideration when formulating a predictive test of articulation.

Hall (1939) reported finding no relationship between good articulation and good sound discrimination, and Hansen (1944) found that speech sound discrimination did not differ between normal and defective adult articulators.

Conversely, Travis and Rasmus (1931) administered a test of 366 pairs of speech sounds to 383 normal articulators and 165 functionally defectives ranging in age from five years to adulthood. They found that at every age level the defectives made significantly more errors than normals, and that a high percentage of sounds missed in the test were those with which these cases were having difficulty. Later, Schiefelbusch and Lindsey (1958) found significant differences between speech defective and normal children in relation to sound discrimination. They (Schiefelbusch and Lindsey, 1958) also found that second-grade normals had better sound discrimination than first-grade normals; but, this was not true of first and second grade defectives. Further evidence in support of the relationship between speech sound discrimination and articulation has been reported by Reid (1947a) and Kronvall and Diehl (1954). Farquhar's (1961) subjects with mild articulatory defects were found to have significantly greater abilities to discriminate sounds among vowels and accoustically dissimilar sounds than did subjects with severe defects. Unfortunately, she found that the auditory discrimination tests designed for her study were not prognostically significant.

More recently, Sherman and Geith (1967) administered the Templin 50-item speech sound discrimination test consisting of 50 pairs of nonsense-syllables to 529 kindergarten children, and selected 18 children with high scores and 18 children with low scores. The Templin-Darley Picture Articulation Test was administered to each of these 36 children. It was found that the high scoring children responded significantly better than the low scorers. Similarly, Marquardt and Saxman (1972) found that children with underdeveloped articulatory skills showed depressed scores in auditory sound discrimination.

There appears to be sufficient evidence to indicate that articulatory skill is related to the level of speech sound discrimination. Weiner (1967) has observed that evidence does support the hypothesis of a link between auditory discrimination and articulation defects. He further observed that this relationship seems to hold in the primary grade age group, until about eight or nine years of age. Similarly, Templin (1943) found that in children from grades two through six, speech sound discrimination improved as grade increased, except for the sixth grade.

Van Riper (1972, pp. 210-213) has been in favor of sound discrimination training in articulation therapy, and Spriesterbach and Curtis (1951) have suggested that ear training be adapted to the particular phonetic contexts in which the individual's misarticulations occur. Locke (1968) has presented theories of sound discrimination and their adaption to therapy from a somewhat different viewpoint. He has suggested the possibility of poor sound discrimination being the result of defective articulatory development. He also pointed out that it is not known whether poor sound discrimination is a cause or a result of defective articulatory performance, but that these theories are in need of further research. It is also possible that traditional sound discrimination tests involving words, syllables, and sounds in isolation are insensitive to whole meanings as they appear in spoken language.

More definite information is necessary before concluding whether or not a relationship between articulatory skill and speech sound discrimination exists, or exactly what that relationship is. However, many of the studies reported here seem to indicate that there is a relationship. Therefore, it may be a mistake not to employ the concept of speech sound discrimination into the formulation of a predictive test of articulation. Sound discrimination and its involvement in the Predictive Screening Test of Articulation will be discussed further in Chapter III, PSTA Rationale.

## Other Predictive Measures of Articulation

This review has been concerned with the major variables related to the prediction of articulatory development. These major variables have been (1) the child's ability to correct his misarticulations after auditory or combined auditory-visual stimulation (imitation), and (2) the child's skills in speech sound discrimination. It may be advantageous to make brief comments on variables which some of the research has shown not to be related to articulatory improvement.

Research by Reid (1947a) showed no relationships between articulatory improvement and auditory memory span, degree of kinaesthetic sensitivity, mental age, intelligence (when intelligence quotient is above 70), personal and social adjustment, or chronological age. Reid's conclusion that articulatory improvement and chronological age are not related is questionable, since much research has indicated that number of misarticulations decreases as children grow older. Pettit (1957) investigated the efficiency of a battery of tests for predicting the articulatory development of five-year-old children. He administered the following battery to each child: test with pure tone audiometer, test of speech perception, test of imitation of non-English sounds, test of imitation and articulation of English sounds, test of memory span, test of gross motor control, test of specialization of movement, test of speed of muscle movement, the California Test of Mental Maturity, and the California Test of

Personality. Pettit found that none of the scores gathered on any of these tests were efficient in predicting articulatory development. However, his results with the test of imitation and articulation of English sounds were in disagreement with the majority of research in that area.

Attempts at the construction of predictive measures have been based on many of the aspects surrounding the concepts involved in speech sound acquisition. Some of these predictive tests will be discussed in the paragraphs which follow.

Arey (1938) described A Diagnostic Profile of the Speech of Children in Grades 1, 2, and 3, which was based on two principle examinable functions: (1) reproductive speech, and (2) free speech. Results obtained from this profile yielded information regarding:

a. The child's speech efficiency level in relation to his own group and to the average achievement level for his age and grade.

b. Specific speech habits which are inefficient or faulty, with an indication of the type and extent of the difficulty.

c. Causitive factors as found in background, health and intelligence, and the condition of the peripheral speech structures.

Standardization of the profile had not been completed at the time this article was published, and no known information is available concerning this profile's levels of validity or reliability.

Research by Carter and Buck (1958) has been suggestive of the usefulness of nonsense-syllable testing as a prognostic tool. They administered spontaneous, imitative, and nonsense-syllable tests to 175 first-grade children with functional articulatory defects. At the beginning of the school year, one group of children was given remedial speech therapy over the nine-month school year, and the other group received no therapy. At the end of the school year, both groups were readministered the spontaneous test. Carter and Buck found that in the group which received no therapy, 71.7 percent of those who achieved 25 percent or more correction on the nonsense-syllable test made 100 percent correction on the final spontaneous test. They concluded that children who were able to correct 75 percent of their misarticulations on the nonsense-syllable test, should be excluded from therapy until the end of the second grade, at which time they should be re-tested and offered therapy if necessary. No data are available from this study to indicate the validity or reliability of the Carter and Buck method.

Nichols (1964), has pointed out that time may be an important drawback involved in using the Carter and Buck method. He cited research in which Sommers, et al. (1961) found that the Carter and Buck method required, on the average, 45 minutes to administer to each subject. Nichols failed to mention, however, that Sommers and his associates found that the Carter and Buck test was an excellent predictor of the improvement in articulation of 25 of their subjects with severe articulation problems. However, the predictive validity of the method was found to be only +.56 (rho), which is a very low value. The rho value is obtained when the non-parametric Spearman Rank-Order Correction Coefficient formula is used. Since this formula is used for the treatment of ranked data, results are only generalizable to the study sample specific to the investigation, and not to general populations. If time was not such an important factor in identifying children for therapy caseloads, and if the validity value were higher and generalizable to larger populations, then the Carter and Buck method might be used as a fairly valid instrument.

Steer and Drexler (1960) have suggested two means of caseload selection. In their study, 93 of the subjects from an earlier investigation by Wilson (1954) were administered the articulation test used by Wilson, and the Templin Non-Diagnostic Word Articulation Test (Templin, 1947b). Steer and Drexler used the data collected by Wilson from the Goodenough Draw A Man Test (Goodenough, 1926), the Vineland Social Maturity Scale, and the articulation test which Wilson used when these children were in kindergarten. Product-moment coefficients of correlation were computed to determine any significant relationships between Wilson's earlier data and the Steer and Drexler data. Significant correlations were not obtained from the Goodenough Draw A Man Test or the Vineland Social Maturity Scale. However, from the coefficients obtained from the articulation tests, Steer and Drexler found that the phonemes, /f, l, θ/, were significant predictive variables. Using these predictive variables, they designed a formula for estimating how many errors a child will have left after articulatory improvement through maturation has occurred: $X_t = .777X_{f,l} + .452X_\theta + .385$, where $X_\theta$ is the predicted Templin score at the sixth grade level, and the $X_{f,l}$ and $X_\theta$ are the total errors in all three positions made by the child in kindergarten on the phonemes indicated. Correlation coefficients obtained with this formula were .504 for Steer and Drexler's control group, and .470 for their experimental group which had received speech improvement. Although these correlation coefficients were significantly different from zero at the .01 level of confidence, they remain drastically low and lack clinical significance. At best, the amounts of commonality present between the variables existent in this method are 25.5 percent for the control group, and 22.1 percent for the experimental group. In other words, only one fourth of the variance involved in the use of this formula is accounted for. These are too low values on which to base any predictive validity.

The second prognostic instrument suggested by Steer and Drexler (1960) involved articulation testing at the beginning and end of the kindergarten year. They suggested that if a child improved during kindergarten, it was likely that he would improve in the future grades. Partial correlation coefficients of -.379 for the speech improvement group, and -.509

for the no speech improvement group were obtained. Again, these correlations reached statistical significance. However, they are low values on which to base any predictive validity. Furthermore, the findings of Steer and Drexler were based on ex post facto research. Kurlinger (pp. 359-360) has mentioned that in this kind of investigation, the researchers are not able to control certain independent variables involved. Here, the chances are good that results are then contaminated. At best, the methods for selecting caseloads suggested by Steer and Drexler, and Carter and Buck, should only be used with caution in a battery of prognostic tests. Unfortunately, such a battery of tests would require much more time than is reasonably available to the typical speech clinician.

Farquhar (1961) investigated the prognostic value of imitative and auditory discrimination tests. She administered imitative articulation tests in isolation, nonsense-syllables, and words, and discrimination tests among vowels, and among similar and dissimilar sounds to two groups of articulatory defective children--a mild and a severe group. She found that the imitative tests were prognostically significant; however, the discrimination tests were not. Farquhar did not suggest cut-off scores or other guidelines for exclusion of children from therapy on the basis of her findings. However, her results did coincide with the majority of the literature which has suggested the prognostic value of imitative testing.

Van Riper (1972, p. 196) has discussed the Laradon Articulation Scale briefly; but, he mentioned that no data are available to indicate its reliability or predictive validity.

The problem of caseload selection remains. The Predictive Screening Test of Articulation (PSTA) will be discussed in the following section. The PSTA may be the best approach to solving the problem which has been introduced to date.

# CHAPTER III

## PSTA RATIONALE

The PSTA is a 47-item, imitative test. The examiner simply follows instructions for the administration of each section (nine in all), and elicits verbal responses from the child on all but two items (Nos. 46 and 47). Each item is scored either correct or incorrect, and the complete test usually requires no longer than ten minutes for administration per child.

The PSTA was designed by Charles Van Riper and Robert Erickson at Western Michigan University. They proposed to develop a measure which could differentiate first grade children who would correct their misarticulations through the process of maturation, from those who would continue to have errors unless remedial speech therapy were provided.

Van Riper and Erickson (1968a) demonstrated an inter-examiner reliability level of .895 by conducting a cross-validation examination. The project examiner (a speech clinician specially trained in the administration of the test) tested one group of first-grade children with functional articulation errors, and other clinicians tested another group of similar children. A product-moment correlation coefficient of .81 was obtained on two randomly selected halves of the PSTA. The reliability coefficient by means of the Spearman-Brown formula was .895. It is noted that these values are not as high as investigators would like them to be; however, they are an indication that the PSTA may have greater predictive usefulness than other predictive measures discussed previously.

The PSTA does involve the predictive variables which are believed to be related to prediction. As mentioned earlier, the PSTA is an imitative test. The examiner provides auditory and visual stimulation (i.e. the child is able to observe the oral structures of the examiner while the test words are spoken) while testing in a face-to-face situation. This rationale in the test's construction is consistent with the findings of Snow and Milisen (1954a,b), Carter and Buck (1958), Farquhar (1961), Smith and Ainsworth (1967), Scott and Milisen (1954b), and others.

The PSTA does not involve a measure of speech sound discrimination in the traditional sense in which similar and dissimilar words or syllables (minimal pairs) are presented. Item 46 approaches this kind of testing in which the child must distinguish between a correct and incorrect presentation of the word "finger." However, this item appears to be more dialect oriented. It seems reasonable to assume that sound discrimination is involved in each item since a child must be able to hear a sound and distinguish it from others before he can use it correctly in a given context. Involvement of the concept of sound discrimination in the predictive test is consistent with the findings of Travis and Rasmus (1931), Schiefelbush and Lindsey (1958), Reid (1947a), Kronvall and Diehl (1954), Farquhar (1961), Sherman and Geith (1967), Templin (1943), and others.

The PSTA also involves inconsistency of misarticulations which is believed to be related to prediction by Templin and Darley (1969), Spriestersbach and Curtis (1951), Van Riper and Irwin (1958, pp. 147-149), McDonald (1964), and others. Specific sounds are presented in various contexts which yields an indirect measure of inconsistency.

Finally, the PSTA yields brief information concerning the motor skills of the child which may be related to prediction (Item 47). Dickson (1962) found that children who retained functional articulatory errors did more poorly on the Oseretsky Tests of Motor Proficiency than did children who mastered their errors through maturation.

Measures of prediction, including the PSTA, inevitably produce two kinds of errors. These are false-positive and false-negative errors. Regarding the PSTA, the former occurs when it is predicted that a child will need speech therapy to master his misarticulations, but will do so without it. The latter occurs when it is predicted that a child will not need speech therapy to correct his misarticulations, when in reality he will. At the suggested cut-off score of 34, Van Riper and Erickson's (1968a) samples of first-grade children from their cross-validation study with functional articulatory errors would have included 54 percent, or 158 of those children in therapy. This cut-off score would have allowed a 37 percent false-positive error margin, and a 30 percent false-negative error margin. This means that of the children (N=158) selected for therapy, 37 percent, or 58 children might possibly have corrected their errors without it, and of the children excluded from therapy, 30 percent, or 41 children might possibly have needed therapy, but would not have received it.

In a recent investigation, Arkebauer and Ohlman (1972) studied the amount of agreement between the PSTA and the Screening Deep Test of Articulation (SDTA) as predictive instruments. They suggested the possibility of raising the cut-off score of the PSTA to as high as 40. This would decrease the false-negative error margin and increase the false-positive error margin. They further suggested that children who score below 40, but who score above another, perhaps 30, also be given the SDTA. Arkebauer and Ohlman have been led to believe that obtained results of the SDTA would serve to provide more specific information as a basis for decisions in selecting children for caseloads. If one accepts their assumption, then the method described is worthy of further research.

It was Van Riper and Erickson's purpose in designing the PSTA to identify or predict children who would require speech correction to master their articulatory errors. As a result of their investigations, they proposed that children who scored below the suggested cut-off score of 34 should be classified as positive subjects, and be included in therapy. Those children who scored 34 or above should be regarded as negative subjects, and should not be included in therapy. This cut-off score separated first-grade children into two groups--positive subjects, or those children to be included in the speech therapy program, and negative subjects, those children to be excluded from speech programs. The problems involved in the clinician's reliance upon the cut-off score alone in caseload selection have been discussed earlier under PSTA Rationale (c.f., pp. 25-26).

## Purpose

It was the purpose of the present investigation and the intended item analyses to describe a method by which children falling into the two main Van Riper and Erickson groups (positives and negatives) can be further classified into the false-positive and false-negative groups at the time of testing, while children are in the first grade. If children were thus identified, it would allow the clinician to make more precise judgments as to which children to include in therapy.

The following hypotheses were posed to substantiate any additions to the present PSTA scoring procedures:

Statement of Hypotheses

1. The false-positive subjects missed more items significantly less frequently than the true-positive subjects than would be expected by chance.
2. The false-negative subjects missed more items significantly more frequently than the true-negative subjects than would be expected by chance.
3. The true-positive subjects missed each of the 47 items significantly more frequently than the true-negative subjects.

## CHAPTER IV

## PROCEDURES

In the sections to follow, background information concerning the earlier studies of Van Riper (1966) and Van Riper and Erickson (1968a, 1968b) has been presented for purposes of clarification. The actual responses of subjects as recorded by Van Riper and Erickson were analyzed in the present investigation. The item analyses were performed over these raw data.

Subjects.--Subjects selected for Van Riper and Erickson's (1966) experimental item pool study were 167 first-grade children from southwestern Michigan. They were judged by state-certified public school speech clinicians as having functionally defective articulation which would warrant their enrollment in speech therapy. In the fall of 1962, the project examiner administered the original PSTA consisting of 111 items to each child. None of the children received therapy during the next two years. In the fall of 1963, children still available for study (N=137) were re-examined by the project examiner by means of a simple phonetic inventory and by the elicitation of samples of spontaneous connected speech. At that time, each subject was classified as still defective or normal articulator. The same procedure was performed in the fall of 1964.

To identify the most predictive of the original 111 items, Van Riper and Erickson performed item analyses over each item to differentiate those children who had acquired normal articulation within two years from those who had not. Of the 57 items thus identified, 47 which required no special materials for their administration were selected for inclusion in the current PSTA. Data from subjects still available for re-examination in 1964 were used in the present investigation.

The subjects selected for Van Riper and Erickson's (1968) cross-validation study consisted of two groups. The first group consisted of 180 first-grade children from Calhoun and Shiawassee Counties in Michigan. These children were administered the PSTA in the first grade and re-examined in the third grade by the project examiner. The second group consisted of 113 children from Tuscola County, Michigan. They were administered the PSTA and re-examined by clinicians other than the project examiner. All of the children used for the study were judged by state-certified speech clinicians as having functional misarticulations which would warrant their enrollment in speech therapy, and no subject had a known clinically significant hearing loss. None were given speech therapy from the fall of 1965 through the fall of 1967. Again these subjects were classified into the still defective and normal groups. Data from children (Group I=144, Group II=81) who were still available for re-examination in 1967 were used in this investigation.

Treatment of Data.--Raw data from Van Riper (1966) and Van Riper and Erickson's (1968b) earlier studies were analyzed in this investigation. From those data, subjects were classified into four study groups. Group A, the true-positive subjects, consisted of children who scored below 34 in the first grade and were judged as still having articulatory defects in the third grade. Group B, the false-positive subjects, consisted of children who fell below the cut-off score in the first grade and were judged as normal articulators in the third grade. Group C, the false-negative subjects, consisted of children who performed as well as or better than the cut-off score in the first grade and were judged as still having articulatory defects in the third grade. Finally, Group D, the true-negative subjects, consisted of children who performed as well as or better than the cut-off score in the first grade and were judged as normal articulators in the third grade.

Sums of negative responses to each item were recorded for each of the four groups--A (N=148), B (N=49), C (N=61), and D (N=101). Those sums were converted into percentage values for each group on all 47 items.

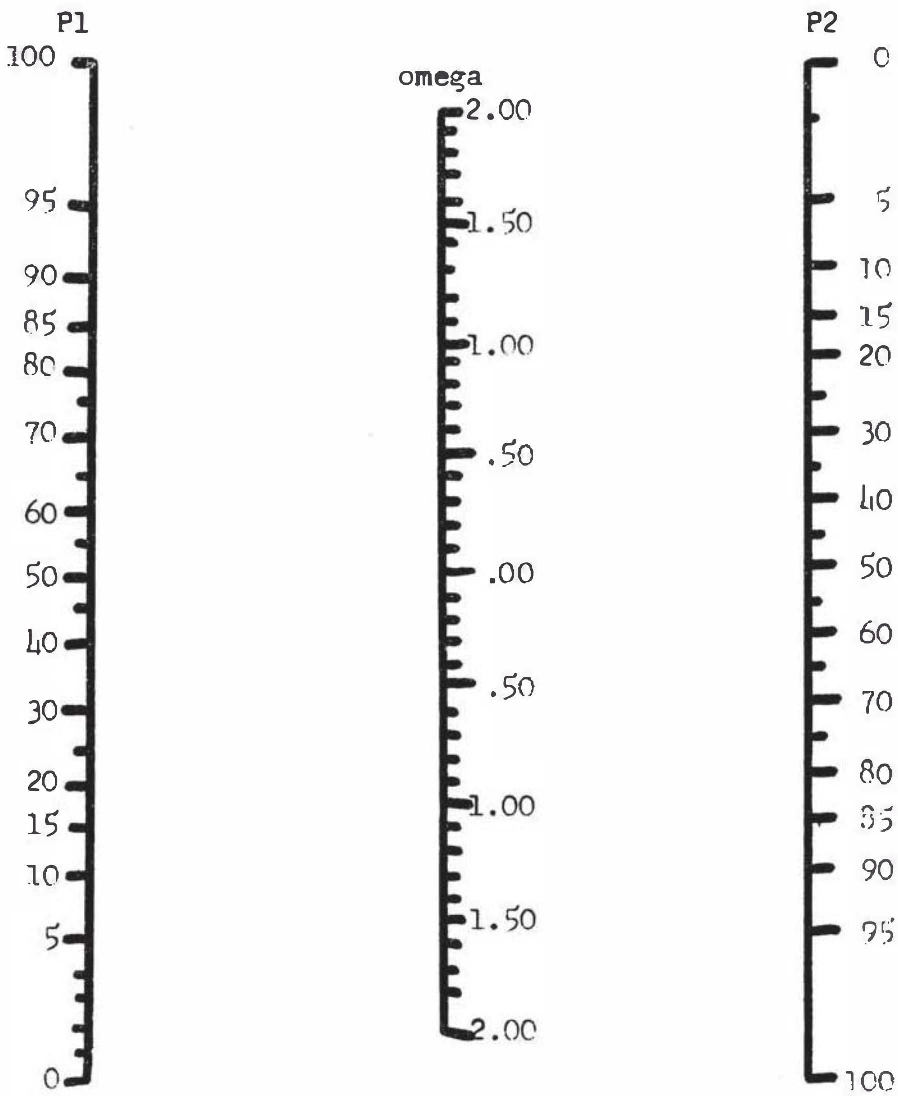
The Lawshe-Baker Nomograph, a test for determining the significance of the difference between two percentage values was used to perform item analyses over each of the

47 items. For purposes of clarification, a reproduction of the Lawshe-Baker Nomograph and the procedures used in its operation has been included in this section (See Figure 4, p. 33).

The purpose of these analyses was to compare and determine significant differences in the frequency of negative responses between the true-positive group (A) and the false-positive group (B), between the false-negative (C) and the true-negative (D) groups, and the true-positive (A) and the true-negative (D) groups. The between-group comparisons are illustrated below:

| Item | A<br>True-<br>Positive | B<br>False-<br>Positive | | C<br>False-<br>Negative | D<br>True-<br>Negative | | A<br>True-<br>Positive | D<br>True-<br>Negative |
|---|---|---|---|---|---|---|---|---|
| 1 | __% | __% | | | | | | |
| 2 | | | | | | | | |
| ↓ | | | | | | | | |
| 47 | | | | | | | | |

FIGURE 4




Lawshe-Baker Nomograph for testing the significance of the difference between two percentages. Two ways of determining significance levels:

When N1 ≠ N2:

.01 level of confidence $\dfrac{2.58}{\sqrt{\dfrac{2(N1\ N2)}{N1 + N2}}}$ .05 level of confidence $\dfrac{1.96}{(\text{same})}$

When N1 = N2 $\dfrac{2.58}{\sqrt{N}}$ $\dfrac{1.96}{\sqrt{N}}$

When the appropriate values for the desired confidence intervals have been computed and located on the center scale, locate the percentage values for P1 and P2 on the respective scales. Connect them with a straight-edge. If the line falls beyond the confidence interval point on the center scale, the percentages are different beyond chance. (From Downie and Heath, 1970, pp. 192-193).

The cut-off score of 34, currently used for interpretation of PSTA results, only yields information which aids in the classification of first-grade children as either positives or negatives. The information gathered by performing the described item analyses should aid the clinician in the classification of first-grade children as false-positives, or false-negatives.

It was expected that for the majority of the 47 items, that comparisons between groups A and D would be significant, since subjects in group A scored below 34, and subjects in group D scored 34 or above. Items which were found not to be significant, may subsequently be regarded as indiscriminative in the classification of children into groups A and D. This would have yielded evidence for the exclusion of these items from the current PSTA, as a result of findings of some further study which exceeds the limits of this present investigation.

# CHAPTER V

# RESULTS AND DISCUSSION

## Results

The purpose of this investigation was to test the following hypotheses:

1. The false-positive subjects missed more items significantly less frequently than the true-positive subjects than would be expected by chance.
2. The false-negative subjects missed more items significantly more frequently than the true-negative subjects than would be expected by chance.
3. The true-positive subjects missed each of the 47 items significantly more frequently than the true-negative subjects.

The purpose of this chapter is to report the results of the statistical analyses performed with the Lawshe-Baker Nomograph as they pertain to each of the three separate hypotheses, and to discuss the implications and the interpretations as they apply to individual comparisons.

TABLE 1

PERCENTAGES OF INCORRECT RESPONSES OF GROUPS A AND B TO THE 47 PSTA TEST ITEMS. RESULTS OF BETWEEN COMPARISONS, OMEGA=.21

| Item | Group A True-Positives | Group B False-Positives | Significant |
|---|---|---|---|
| 1 | .40 | .20 | * |
| 2 | .66 | .51 | |
| 3 | .22 | .03 | * |
| 4 | .66 | .62 | |
| 5 | .67 | .62 | |
| 6 | .33 | .18 | |
| 7 | .55 | .43 | |
| 8 | .62 | .48 | |
| 9 | .31 | .20 | |
| 10 | .49 | .38 | |
| 11 | .32 | .20 | |
| 12 | .35 | .21 | |
| 13 | .23 | .15 | |
| 14 | .24 | .20 | |
| 15 | .22 | .12 | |
| 16 | .13 | .07 | |
| 17 | .18 | .07 | * |
| 18 | .39 | .21 | * |
| 19 | .39 | .23 | |
| 20 | .33 | .15 | * |
| 21 | .35 | .16 | * |
| 22 | .45 | .23 | * |
| 23 | .51 | .28 | * |
| 24 | .30 | .12 | * |
| 25 | .35 | .12 | * |
| 26 | .74 | .87 | |
| 27 | .74 | .77 | |
| 28 | .74 | .79 | |
| 29 | .72 | .75 | |
| 30 | .74 | .79 | |
| 31 | .74 | .80 | |
| 32 | .28 | .12 | * |
| 33 | .71 | .54 | * |
| 34 | .29 | .15 | |
| 35 | .29 | .13 | * |
| 36 | .74 | .75 | |
| 37 | .80 | .89 | |
| 38 | .80 | .84 | |
| 39 | .95 | .97 | |
| 40 | .57 | .53 | |
| 41 | .27 | .15 | |
| 42 | .60 | .57 | |
| 43 | .64 | .61 | |
| 44 | .39 | .43 | |
| 45 | .20 | .12 | |
| 46 | .17 | .12 | |
| 47 | .38 | .21 | * |

*Significant at or beyond the .05 level of confidence

TABLE 2

PERCENTAGES OF INCORRECT RESPONSES OF GROUPS C AND D TO THE 47 PSTA TEST ITEMS. RESULTS OF BETWEEN COMPARISONS, OMEGA=.25

| Item | Group C False-Negatives | Group D True-Negatives | Significant |
|---|---|---|---|
| 1 | .25 | .17 | |
| 2 | .16 | .10 | |
| 3 | .08 | .03 | |
| 4 | .27 | .14 | |
| 5 | .10 | .14 | |
| 6 | .12 | .10 | |
| 7 | .45 | .31 | |
| 8 | .51 | .40 | |
| 9 | .25 | .16 | |
| 10 | .49 | .32 | |
| 11 | .16 | .05 | * |
| 12 | .18 | .05 | * |
| 13 | .14 | .04 | * |
| 14 | .06 | .03 | |
| 15 | .08 | .00 | * |
| 16 | .10 | .01 | * |
| 17 | .08 | .00 | * |
| 18 | .10 | .08 | |
| 19 | .20 | .11 | |
| 20 | .16 | .10 | |
| 21 | .20 | .09 | |
| 22 | .24 | .15 | |
| 23 | .33 | .31 | |
| 24 | .04 | .05 | |
| 25 | .02 | .07 | |
| 26 | .10 | .08 | |
| 27 | .06 | .08 | |
| 28 | .08 | .08 | |
| 29 | .08 | .08 | |
| 30 | .08 | .09 | |
| 31 | .12 | .15 | |
| 32 | .00 | .03 | |
| 33 | .55 | .45 | |
| 34 | .14 | .09 | |
| 35 | .12 | .06 | |
| 36 | .14 | .14 | |
| 37 | .20 | .18 | |
| 38 | .24 | .16 | |
| 39 | .59 | .45 | |
| 40 | .24 | .03 | * |
| 41 | .37 | .26 | |
| 42 | .14 | .08 | |
| 43 | .20 | .01 | * |
| 44 | .39 | .35 | |
| 45 | .08 | .04 | |
| 46 | .14 | .07 | |
| 47 | .33 | .24 | |

*Significant at or beyond the .05 level of confidence

TABLE 3

PERCENTAGES OF INCORRECT RESPONSES OF GROUPS A AND D TO THE 47 PSTA TEST ITEMS. RESULTS OF BETWEEN COMPARISONS, OMEGA=.18

| Item | Group A True-Positives | Group D True-Negatives | Significant |
|---|---|---|---|
| 1 | .40 | .17 | * |
| 2 | .66 | .10 | * |
| 3 | .22 | .03 | * |
| 4 | .66 | .14 | * |
| 5 | .67 | .14 | * |
| 6 | .33 | .10 | * |
| 7 | .55 | .31 | * |
| 8 | .62 | .40 | * |
| 9 | .31 | .16 | * |
| 10 | .49 | .32 | * |
| 11 | .32 | .05 | * |
| 12 | .35 | .05 | * |
| 13 | .23 | .04 | * |
| 14 | .24 | .03 | * |
| 15 | .22 | .00 | * |
| 16 | .13 | .01 | * |
| 17 | .18 | .00 | * |
| 18 | .39 | .08 | * |
| 19 | .39 | .11 | * |
| 20 | .33 | .10 | * |
| 21 | .35 | .09 | * |
| 22 | .45 | .15 | * |
| 23 | .51 | .31 | * |
| 24 | .30 | .05 | * |
| 25 | .35 | .07 | * |
| 26 | .74 | .08 | * |
| 27 | .74 | .08 | * |
| 28 | .74 | .08 | * |
| 29 | .72 | .08 | * |
| 30 | .74 | .09 | * |
| 31 | .74 | .15 | * |
| 32 | .28 | .03 | * |
| 33 | .71 | .45 | * |
| 34 | .29 | .09 | * |
| 35 | .29 | .06 | * |
| 36 | .74 | .14 | * |
| 37 | .80 | .18 | * |
| 38 | .80 | .16 | * |
| 39 | .95 | .45 | * |
| 40 | .57 | .03 | * |
| 41 | .27 | .26 | |
| 42 | .60 | .08 | * |
| 43 | .64 | .01 | * |
| 44 | .39 | .35 | |
| 45 | .20 | .04 | * |
| 46 | .17 | .07 | * |
| 47 | .38 | .24 | |

*Significant at or beyond the .05 level of confidence

Hypothesis No. 1.--The results obtained from this investigation supported hypothesis No. 1. As shown in Table 1, p. 36, comparisons made between the incorrect responses of groups A and B revealed that items 1, 3, 17, 18, 20, 21, 22, 23, 24, 25, 32, 35, and 47 were statistically different at or beyond the .05 level of confidence. That is, group B responded incorrectly to these items significantly less frequently than did group A.

Hypothesis No. 2.--The results obtained support hypothesis No. 2. As shown in Table 2, p. 37, comparisons made between the incorrect responses of groups C and D revealed that items 11, 12, 13, 15, 16, 17, 40, and 43 were significantly different at or beyond the .05 level of confidence. That is, group C responded incorrectly to these items significantly more frequently than did group D.

Hypothesis No. 3.--Results of comparisons between the incorrect responses of the true-positive (A) and the true-negative (D) groups fail to support hypothesis No. 3. As shown in Table 3, p. 38, between comparisons were significantly different at or beyond the .05 level of confidence for all but three items--Nos. 41, 44, and 47. That is, group A responded incorrectly to all but three of the test items significantly more frequently than did group D.

## Discussion

Hypothesis No. 1.--As a result of comparisons between the incorrect responses of groups A and B to each item, it was found that the false-positive group (B) responded incorrectly to items 1, 3, 17, 18, 20, 21, 22, 23, 24, 25, 32, 33, 35, and 47 significantly less frequently than did the true-positive group (A). By referring to Table 3, p. 38, the reader should notice that all of these items except for No. 47 were found to be significantly different when the responses of the true-positive group (A) and the true-negative group (D) were compared. This is interpreted to mean that all items presented above, except for No. 47, are truly discriminating test items for the purposes of classifying children into groups A and D. All items except No. 47 satisfy two criteria: (1) they appear to discriminate between responses of the true-positive and the false-positive groups; and, (2) they appear to discriminate between the responses of the true-positive and true-negative groups. It is suggested that because item 47 did not satisfy the second criterion, that it should not be considered to be a discriminating item in the classification of children into any of the four groups. It is possible that for the A-D comparison of responses to item 47, that the difference may be attributed to chance. Further investigation would be necessary to determine the variables influencing subject performance on item 47.

Hypothesis No. 2.--Comparisons between the incorrect responses of groups C and D revealed that the false-negative group (C) responded incorrectly significantly more frequently to items 11, 12, 13, 15, 16, 17, 40, and 43 than did the true-negative group (D). This is interpreted to mean that the items presented above were truly discriminating items in the classification of children into groups C and D. Each of these items satisfy two criteria: (1) they appear to discriminate between the responses of the false-negative and true-negative groups.

Hypothesis No. 3.--As a result of comparisons between the incorrect responses of groups A and D, it was found that the true-positive group responded incorrectly significantly more frequently than did the true-negative group to all but three items--Nos. 41, 44, and 47. This is interpreted to mean that items 41, 44, and 47 were not discriminating items in the classification of children into the major study groups A and D, the positives and the negatives. That is, items 41, 44, and 47 do not truly aid the clinician in separating children who actually need therapy from those children who do not.

As a result of this investigation, these conclusions seem warranted concerning the predictive validity of the PSTA:

1. The suggested cut-off score (34) initially separates children into either of two groups--the positives (those requiring speech therapy) and the negatives (those not

requiring therapy).

2. The suggested cut-off score allows for type I and II errors. That is, children are classified as requiring speech therapy when they actually do not, and children are classified as not requiring therapy when they actually do.

3. The type I errors, or the false-positive subjects, appear to respond incorrectly significantly less frequently to specific test items than do the true-positive subjects.

4. The type II errors, or the false-negative subjects, appear to respond incorrectly significantly more frequently to specific test items than do true-negative subjects.

5. Three of the items included in the current PSTA do not appear to aid in the successful classification of subjects into the positive and negative groups. That is, these items do not appear to discriminate between children who require therapy from children who do not.

It was not the purpose of this investigation to obtain evidence for the inclusion or the exclusion of current test items. The original designers of the PSTA may elect to exclude items 41, 44, and 47 if they find it useful to do so. It is unlikely that the inclusion of these items causes spurious classification of subjects. If the authors should elect to exclude these items, then it shall become necessary that a new cut-off score be computed relative to the total number of test items.

It was the purpose of this investigation, however, to identify items for further classification of children into the false-positive and false-negative groups. This purpose was accomplished. A total of 141 comparisons were made in this investigation. In consideration of that large a number, it is likely that seven of those comparisons may have differed significantly due to chance. However, the number of significant comparisons far exceeded that number which could be attributed to chance. This is interpreted to mean that obtained differences were real differences. These real differences in between comparisons serve to enlarge the usefulness of these data considerably.

It was found that the false-positive subjects missed items 1, 3, 17, 18, 20, 21, 22, 23, 24, 25, 32, 33, and 35 significantly less frequently than did the true-positive subjects. In reference to the Appendix of this report, it is observed that all of these items involve the glide consonants /r/ (Tiffany, 1960, p. 27; Hanley and Thurman, 1970, pp. 90-91) and /l/ (Kantner and West, 1960, pp. 297-299) appearing in words as single elements and in consonant combinations, except for item 17, which involves the affricate /dʒ/. Item 17 was the only item found to be significant for both comparisons between groups A and B, and groups C and D. It is suggested that the significant comparison between groups A and B on item 17 be regarded as a sampling error since that item seems to be consistent with the other items found to be significant between the groups C and D comparisons.

For the comparisons between groups C and D, it was found that items 11, 12, 13, 15, 17, 40, and 43 were significant ($p < .05$). That is, the false-negative subjects missed these items significantly more frequently than did the true-negative subjects. Again, in reference to the Appendix, p.56, it is observed that these items involve the later developed fricative cognates /s/ and /z/, the fricative /ʃ/, and the affricate cognates /tʃ/ and /dʒ/.

It appears that the false-positive and the false-negative subjects did, indeed, miss specific kinds of items which would serve to classify them into groups other than the positives and the negatives. The manner of articulation (Hanley and Thurman, 1970, p. 83) appears to be the discriminating articulatory element between the items which distinguish the false-positives from the true-positives, and the false-negatives from the true-negatives. With respect to the distinctive feature theory set forth by Chomsky and Halle (1968, p. 177), the features which appear to distinguish the /r/ and the /l/ from the /s/, /z/, /tʃ/, /dʒ/ and the /ʃ/, are the rules which apply to sounds which are vocalic and strident (Table 4, p. 45). Vocalic sounds are produced with an oral cavity in which the most radical constriction does not exceed that found in the high vowels /i/ and /u/ and with vocal cords that are positioned so as to allow spontaneous voicing (Chomsky and Halle, 1968, p. 302). In reference to the phonemes /r/ and /l/ which appear to distinguish false-positives from true-positives, it would

TABLE 4

DISTINCTIVE FEATURES WHICH APPLY TO PHONEMES /r/, /l/, /s/, /z/, /tʃ/, /dʒ/, and /ʃ/

Phoneme

| Rule | r | l | s | z | tʃ | dʒ | ʃ |
|---|---|---|---|---|---|---|---|
| Vocalic | + | + | - | - | - | - | - |
| Consonantal | + | + | + | + | + | + | + |
| High | - | - | - | - | + | + | + |
| Back | - | - | - | - | - | - | - |
| Low | - | - | - | - | - | - | - |
| Anterior | - | + | - | - | - | - | - |
| Coronal | + | + | + | + | + | + | + |
| Voice | + | + | - | + | - | + | - |
| Continuant | + | + | + | + | - | - | + |
| Nasal | - | - | - | - | - | - | - |
| Strident | - | - | + | + | + | + | + |

seem that the concept of spontaneous vocal cord anticipation of voicing may be a significant factor to consider when articulatory maturation is measured. Strident sounds are marked acoustically by greater noisiness than their non-strident counterparts (Chomsky and Halle, 1968, p. 329). In reference to the phonemes /s/, /z/, /tʃ/, /dʒ/, and /ʃ/, it appears that the acoustic characteristics of these phonemes might bear considerable influence in the distinction between the false-negatives and the true-negatives. It is possible that the true-negative subjects of Van Riper and Erickson's (1968b) earlier studies were significantly more sophisticated in the recognition of the acoustical characteristics of the /s/, /z/, /tʃ/, /dʒ/, and /ʃ/ than were the false-negative subjects. Further research in the area of distinctive features as they relate to the measurement of articulatory maturation seems warranted.

In completion of this investigation, items 1, 3, 18, 20, 21, 22, 23, 24, 25, 32, 33, and 35, which appear to separate the false-positive from the true-positive subjects, have been appended to the original Response Record Sheet (Appendix, p.56). In like manner, items 11, 12, 13, 15, 16, 17, 40, and 43, which appear to separate the false-negative from the true-negative subjects, have also been appended.

In reference to the results of the present investigation, the following are suggested steps for the clinician to employ in the interpretation of PSTA results:

1. Administer and score the PSTA following the instructions as written in the PSTA Manual (Van Riper and Erickson, 1968b, pp. 9-16). Use the Response Record Sheet as shown in the Appendix, p. 56 of this report.

2. After results of tests are obtained, separate children who scored below 34 from those who scored 34 and above. From information obtained by interpretation of the cut-off score, children are classified as positive (below 34) and negative (34 and above) subjects.

3. It is now suggested that children be further classified into false-positive and false-negative groups. First examine the response sheets of children who scored below 34. Refer to the bottom of the Response Record Sheet where "False-Positive" is printed, and the numbered items follow. Circle those items to which the child responded correctly. Follow the same procedure for the response sheets of children who scored 34 and above, except, refer to the items following "False-Negatives," and circle the items to which the child responded incorrectly.

4. Interpretation of these results will depend on the clinician's therapy caseload. Some of the children are likely to make correct responses to all items following "False-Positive," and some are likely to respond incorrectly to all of the items following "False-Negative." In this case, identification of children into the false-positive and false-negative groups should be quite simple. Those

children for whom no items are circled should be regarded as true-positives and true-negatives.

It is likely that the clinician using this approach will find that for a given number of children, the numbers of circled items may vary. The clinician might ask "How many items must be circled to be reasonably assured that the child will fall into either the false-positive or the false-negative group?" A specific answer to that question is not yet at hand. The question should be answered with the results of future research. It is reasonable to assume, however, that as the number of circled items increases, so does the probability of that child falling into a specific group. It is necessary that the clinician use clinical insight when deciding which children to select for therapy. The size of her caseload should influence her judgments concerning the number of items which must be circled for a child to be classified into any given group.

CHAPTER VI

# SUMMARY AND IMPLICATIONS

## Summary

The purpose of this chapter is to summarize the rationale behind the present investigation, the procedures involved, and the results thereof. The investigator has also included a section which expresses implications for further research derived from the present study.

It was the purpose of the present investigation to test the following hypotheses:

1. The false-positive subjects missed more items significantly less frequently than the true-positive subjects than would be expected by chance.
2. The false-negative subjects missed more items significantly more frequently than the true-negative subjects than would be expected by chance.
3. The true-positive subjects missed each of the 47 items significantly more frequently than the true-negative subjects.

It was the investigator's purpose in testing hypothesis

No. 1 to identify specific PSTA items which differentiated false-positive subjects from true-positive subjects. With respect to hypothesis No. 2, the investigator proposed to identify specific items which differentiated the false-negatives from the true-negatives. The reason for testing hypothesis No. 3 was to determine whether or not each of the 47 test items was a truly discriminating item in distinguishing children who would require therapy from those who would not.

The basic problem was stated. That was, the speech clinician had no valid means for deciding which children with functional articulatory errors in the first grade would actually require speech correction to master their articulatory errors. It was pointed out that this was a serious problem to the clinician since research has indicated that from 75 to 80 percent of the children included in speech correction programs consisted of functional articulatory defectives.

A review of the literature indicated that speech sound acquisition seemed to be a gradual process which occurred with maturation. Studies by Louttit and Halls (1936), Mills and Striet (1942), Templin (1957), Poole (1934), Roe and Milisen (1942), Williams (1971), and others showed that misarticulations appeared to decrease with age, until the child reached the third or fourth grades.

Studies by Snow and Milisen (1954b), Carter and Buck (1958), Farquhar (1961), Nichols (1964), and others presented information which indicated that a child's ability to correct misarticulations after auditory stimulation may be a strong indication that spontaneous improvement would occur, and that imitative articulatory testing is likely to be a useful and valid predictive tool.

Sound discrimination and its relationship to articulatory performance was reviewed. Research by Travis and Rasmus (1931), Reid (1947a), Kronvall and Diehl (1954), Sherman and Geith (1967), and others supported the argument for such a relationship. However, conflicting results were reported by Hall (1939), and Hansen (1944) who found no relationship. The majority of the findings in this area led the investigator to conclude that a relationship between sound discrimination and articulatory performance may exist, and that speech sound discrimination might possibly be a useful predictive tool.

Attempts at the construction of other predictive measures were reviewed. Among these were the Diagnostic Profile described by Arey (1938), and methods of predicting articulatory improvement presented by Carter and Buck (1958), and Steer and Drexler (1960). Little information was known concerning the Diagnostic Profile. Sommers, et al. (1961) found that the Carter and Buck method had low predictive validity (rho= .56). Steer and Drexler attempted to design a formula by which a child's score on the Templin Non-Diagnostic Word Articulation Test could be predicted at the sixth grade level. Obtained correlation coefficients were .504 and .470.

These levels were too low to base predictive validity on that formula. Steer and Drexler suggested a second method which involved articulation testing at the beginning and end of the kindergarten year. Obtained correlation coefficients were again low (-.379 and -.509). It was concluded that the Carter and Buck and the Steer and Drexler methods should only be used with caution in a battery of articulation tests.

In support of the Predictive Screening Test of Articulation, the investigator argued that the PSTA contained the predictive variables necessary in the formulation of a valid predictive instrument. Problems with the interpretation of PSTA results were discussed. It was pointed out that the PSTA is capable of allowing type I and type II errors which serve to decrease the test's level of predictive validity. It was the investigator's intention to identify specific test items which would serve to further identify the false-positive and the false-negative subjects at the time of testing.

The procedures involved in the execution of the present investigation were discussed. Raw data from the subjects of Van Riper (1966) and Van Riper and Erickson's (1968a, 1968b) earlier studies were analyzed. Comparisons were made between the incorrect responses assigned to the true-positives and false-positives, between the false-negatives and the true-negatives, and between the true-positives and the true-negatives. The Lawshe-Baker Nomograph was used to perform the statistical analyses. It was expected that the false-positive subjects missed specific items less frequently than the true-positives, and that the false-negative subjects missed specific items more frequently than the true-negatives. It was further expected that the true-positives missed each of the 47 items more frequently than the true-negatives.

The results were presented in order of the stated hypotheses. With respect to hypothesis No. 1, it was found that the false-positives missed items 1, 3, 17, 18, 20, 21, 22, 23, 24, 25, 32, 35, and 47 significantly less frequently than the true-positives. Hence, hypothesis No. 1 was confirmed. Hypothesis No. 2 was also confirmed. The false-negative subjects missed items 11, 12, 13, 15, 17, 40, and 43 significantly more frequently than the true-negatives. Hypothesis No. 3 was not supported when comparisons were made between the incorrect responses of the true-positives and the true-negatives. It was found that items 41, 44, and 47 were non-significant.

Interpretations of these results were presented. With respect to hypothesis No. 1, significant items, except for Nos. 17 and 47, were regarded as truly discriminating items for purposes of distinguishing false-positive from true-positive subjects. Item 17 was considered significant due to sampling error, since it was not consistent with the other items involving the /r/ and the /l/. The significant comparison for item 47 was also regarded as a sampling error since that item was not significant for the comparison between the true-positive and the true-negative subjects.

Item 17 was also a significant item when comparisons were made between the false-negatives and the true-negatives. This item was consistent with the other items involving the /s/, /z/, /tʃ/, /dʒ/, and the /ʃ/. All of the other significant items for the false-negative and true-negative comparisons were regarded as truly discriminating items for purposes of distinguishing the false-negative from the true-negative subjects.

In consideration of hypothesis No. 3, since items 41, 44, and 47 were found to be non-significant, they were regarded as indiscriminative in the distinction between the true-positives and the true-negatives. It was suggested to Van Riper and Erickson that it might be well to delete these items from the current PSTA.

In completion of this report, the words "False-Positive" followed by items numbered 1, 3, 18, 20, 21, 22, 23, 24, 25, 32, 33, and 35, and "False-Negative" followed by items numbered 11, 12, 13, 15, 16, 17, 40, and 43 were appended to the bottom of the original response sheet (Appendix, p.56). It was suggested to the clinician using the PSTA that he refer to these items to identify children who are likely to fall into the false-positive and false-negative groups.

## Implications for Further Research

The following implications for further research were derived from this investigation:

1. The present investigation did not propose to demonstrate evidence for changing actual PSTA items. However, research by Leonard and Ritterman (1971) has indicated that the frequency of occurrence of stimulus words in the language influences articulatory performance. The research of Arkebauer and Ohlman (1972) indicated that the PSTA may overemphasize sibilants and neglect other phonemes, and that consistent misarticulation of other phonemes might also be worthy of identification. Further research seems necessary to test the predictive usefulness of individual PSTA items, and other test stimuli which may have predictive value.

2. Chase (1973) has investigated the agreement among PSTA scores assigned by inexperienced and experienced judges to the responses of three first-grade children to each of the 47 items. His results indicated that some children may be more difficult to score than others. He also found that some of the items currently included in the PSTA appeared more difficult to score than others. These findings imply that the clinician using the PSTA should have some indication as to which and how many children he tests may be scored spuriously. Further research in this area seems justified. It also seems that research is necessary to identify current PSTA items which, in themselves, may be difficult to score.

## APPENDIX

### RESPONSE RECORD SHEET

Record the child's response to each item of the PSTA by circling the 1 if his response is correct or by circling the 0 if his response is incorrect (or if no respnse is made). Compute the child's "Total Score" by counting the number of items where 1 has been circled. Refer to the words "False-Positive" and "False-Negative" at the bottom.. If the child falls below the cut-off score, circle the items he passed following "False-Positive." If he scores 34 or above, circle the items he missed following "False-Negative."

| Item | Response | Item | Response | Item | Response |
|---|---|---|---|---|---|
| Part I | | Part III | | | |
| 1. RABBIT | 1 0 | 18. PRESENTS | 1 0 | 36. SLED | 1 0 |
| 2. SOAP | 1 0 | 19. BREAD | 1 0 | 37. SPLASH | 1 0 |
| 3. LEAF | 1 0 | 20. CRAYONS | 1 0 | 38. STRING | 1 0 |
| 4. ZIPPER | 1 0 | 21. GRASS | 1 0 | Part IV | |
| Part II | | 22. FROG | 1 0 | 39. Sentence | 1 0 |
| 5. MUSIC | 1 0 | 23. THREE | 1 0 | Part V | |
| 6. VALENTINE | 1 0 | 24. CLOWN | 1 0 | 40. /s/ | 1 0 |
| 7. TEETH | 1 0 | 25. FLOWER | 1 0 | 41. /θ/ | 1 0 |
| 8. SMOOTH | 1 0 | 26. SMOKE | 1 0 | Part VI | |
| 9. ARROW | 1 0 | 27. SNAKE | 1 0 | 42. SEESEESEE | 1 0 |
| 10. BATHTUB | 1 0 | 28. SPIDER | 1 0 | 43. ZOOZOOZOO | 1 0 |
| 11. SHEEP | 1 0 | 29. STAIRS | 1 0 | 44. PUHTUHKUH | 1 0 |
| 12. DISHES | 1 0 | 30. SKY | 1 0 | Part VII | |
| 13. CHAIR | 1 0 | 31. SWEEPING | 1 0 | 45. LA-LA-LA | 1 0 |
| 14. MATCHES | 1 0 | 32. PLANT | 1 0 | Part VIII | |
| 15. WATCH | 1 0 | 33. SHREDDED WHEAT | 1 0 | 46. /ɚ/ Recognition | 1 0 |
| 16. JAR | 1 0 | 34. TREE | 1 0 | Part IX | |
| 17. ENGINE | 1 0 | 35. DRESS | 1 0 | 47. Clapping rhythm | 1 0 |

**False-Positive:** 1 3 18 20 21 22 23 24 25 32 33 35

**False-Negative:** 11 12 13 15 16 17 40 43

# BIBLIOGRAPHY

## Books


Carrell, James, and Tiffany, William R. Phonetics: Theory and Application to Speech Improvement. New York: McGraw Hill, 1960.

Chomsky, Noam, and Halle, Morris. The Sound Pattern of English. New York: Harper and Row, 1968.

Downie, N. M., and Heath, R. W. Basic Statistical Methods. 3rd ed. New York: Harper and Row, 1970.

Goodenough, Florence Laura. Measurement of Intelligence by Drawings. New York: World Book Co., 1926.

Hanley, Theodore D., and Thurman, Wayne L. Developing Vocal Skills. 2nd ed. New York: Holt, Rinehart, and Winston, 1970.

Johnson, Wendell; Darley, Frederick L.; and Spriestersbach, D. C. Diagnostic Methods in Speech Pathology. New York: Harper and Row, 1963.

Kantner, Claude E., and West, Robert. Phonetics. New York: Harper and Bros., 1960.

Kurlinger, Fred N. Foundations of Behavioral Research. New York: Holt, Rinehart, and Winston, 1964.

McDonald, Eugene T. Articulation Testing and Treatment: A Sensory Motor Approach. Pittsburgh: Stanwix House, 1964.

Powers, Margaret Hall. "Functional Disorders of Articulation--Symptomatology and Etiology." Handbook of Speech Pathology and Audiology. Edited by Lee Edward Travis. New York: Appleton-Century-Crofts, 1971.

Templin, Mildred C. Certain Language Skills in Children, Their Development and Interrelationships. Minneapolis: University of Minnesota Press, 1957.

Van Riper, Charles. Speech Correction: Principles and Methods. 5th ed. Englewood Cliffs: Prentice-Hall, 1972.

Van Riper, Charles, and Irwin, John V. Voice and Articulation. 3rd ed. Englewood Cliffs: Prentice-Hall, 1961.

## Articles

Arey, Mabel-Louise. "A Diagnostic Profile of the Speech of Children in Grades 1, 2, and 3." Quarterly Journal of Speech, 24 (No. 2, 1938), 265-8.

Bingham, Dale S.; Van Hattum, Rolland J.; Faulk, Margaret E.; and Taussig, Eleanor. "Program Organization and Management." "Public School Speech and Hearing Services." Journal of Speech and Hearing Disorders, (Monograph Supplement 8, 1961), 33-49.

Carter, Eunice T., and Buck, McKensie. "Prognostic Testing for Functional A ticulation Disorders Among Children in the First Grade." Journal of Speech and Hearing Disorders, 23 (No. 2, 1958), 124-133.

Dawson, L. O. "A Study of the Development of the Rate of Articulation." Elementary School Journal, 29 (No. 8, 1929), 610-5.

Dickson, Stanley. "Differences Between Children Who Spontaneously Outgrow and Children Who Retain Functional Articulation Errors." Journal of Speech and Hearing Research, 5 (No. 3, 1962), 263-71.

Farquhar, Mary Stuart. "Prognostic Value of Imitative and Auditory Discrimination Tests." Journal of Speech and Hearing Disorders, 26 (No. 4, 1961), 342-7.

Hall, Margaret E. "Auditory Factors in Functional Articulatory Speech Defects." Journal of Experimental Education, 7 (No. 2, 1938), 110-132.

Hansen, Burrell F. "The Application of Sound Discrimination Tests to Functional Articulatory Defectives with Normal Hearing." Journal of Speech Disorders, 9 (No. 4, 1944) 347-55.

Irwin, Orvis C. "Infant Speech: Consonant Sounds According to Manner of Articulation." Journal of Speech Disorders, 12 (No. 4, 1947), 402-4.

Irwin, Orvis C., and Curry, Thayer. "Vowel Elements in the Crying Vocalizations of Infants Under Ten Days of Age." *Child Development*, 12 (No. 2, 1941), 99-109.

Kronvall, Ernest L., and Diehl, Charles F. "The Relationship of Auditory Discrimination to Articulatory Defects of Children with No Known Organic Impairment." *Journal of Speech and Hearing Disorders*, 19 (No. 3, 1954), 335-8.

Leonard, Laurence B., and Ritterman, Stuart I. "Articulation of /s/ as a Function of Cluster and Word Frequency of Occurrence." *Journal of Speech and Hearing Research*, 14 (No. 3, 1971), 476-85.

Locke, John L. "Discriminative Learning in Children's Acquisition of Phonology." *Journal of Speech and Hearing Research*, 11 (No. 2, 1968), 428-34.

Loutitt, C. M., and Halls, E. C. "Survey of Speech Defects Among Public School Children of Indiana." *Journal of Speech Disorders*, 1 (Sept., 1936), 73-80.

Marquardt, Thomas P., and Saxman, John H. "Language Comprehension and Auditory Discrimination in Articulation Deficient Kindergarten Children." *Journal of Speech and Hearing Research*, 15, (No. 2, 1972), 382-9.

Metraux, Ruth W. "Speech Profiles of the Pre-School Child 18 to 54 Months." *Journal of Speech and Hearing Disorders*, 15 (No. 1, 1950), 37-53.

Mills, Alice W., and Streit, Helen. "Report of a Speech Survey, Holyoke, Massachusetts." *Journal of Speech Disorders*, 7 (No. 2, 1942), 161-7.

Nichols, Alan C. "Allocation of Time in the Articulation Program: Applications of Research." *ASHA*, 6 (No. 1, 1964), 8-12.

Pettit, Calvin W. "The Predictive Efficiency of a Battery of Articulatory Diagnostic Tests." *Speech Monographs*, 24 (No. 3, 1957), 219-226.

Poole, Irene. "Genetic Development of Articulation of Consonant Sounds in Speech." *Elementary English Review*, 11 (No. 6, 1934), 159-61.

Reid, Gladys. "The Efficacy of Speech Re-education of Functional Articulatory Defectives in the Elementary School." *Journal of Speech Disorders*, 12 (No. 3, 1947b), 301-13.

Reid, Gladys. "The Etiology and Nature of Functional Articulatory Defects in Elementary School Children." Journal of Speech Disorders, 12 (No. 2, 1947a), 143-50.

Roe, Vivian, and Milisen, Robert. "The Effect of Maturation Upon Defective Articulation in Elementary Grades." Journal of Speech Disorders, 7 (No. 1, 1942), 37-50.

Root, A. R. "A Survey of Speech Defectives in the Public Elementary Schools of South Dakota." Elementary School Journal, 26 (No. 6, 1926), 531-41.

Sander, Eric K. "When Are Speech Sounds Learned?" Journal of Speech and Hearing Disorders, 37 (No. 1, 1972), 55-63.

Saylor, Helen K. "The Effect of Maturation Upon Defective Articulation in Grades Seven Through Twelve." Journal of Speech and Hearing Disorders, 14 (No. 3, 1949), 202-7.

Schiefelbusch, R. L., and Lindsey, Mary Jeanne. "A New Test of Sound Discrimination." Journal of Speech and Hearing Disorders, 23 (No. 2, 1958), 153-9.

Scott, Davis A., and Milisen, Robert. "The Effect of Visual, Auditory and Combined Visual-Auditory Stimulation Upon the Speech Responses of Defective Speaking Children." Journal of Speech and Hearing Disorders, (Monograph Supplement 4, 1954a), 38-43.

Scott, Davis A., and Milisen, Robert. "The Effectiveness of Combined Visual-Auditory Stimulation in Improving Articulation." Journal of Speech and Hearing Disorders, (Monograph Supplement 4, 1954b), 52-6.

Sherman, Dorothy, and Geith, Annette. "Speech Sound Discrimination and Articulation Skill." Journal of Speech and Hearing Research, 10 (No. 2, 1967), 277-80.

Siegel, Gerald; Winitz, Harris; and Conkey, Harlan. "The Influence of Testing Instrument on Articulatory Responses of Children." Journal of Speech and Hearing Disorders, 28 (No. 1, 1963), 67-76.

Smith, Myrna W., and Ainsworth, Stanley. "The Effects of Three Types of Stimulation on Articulatory Responses of Speech Defective Children." Journal of Speech and Hearing Research, 10 (No. 2, 1967), 333-8.

Snow, Katherine, and Milisen, Robert. "The Influence of Oral Versus Pictorial Presentation Upon Articulation Testing Results." Journal of Speech and Hearing Disorders, (Monograph Supplement 4, 1954a), 29-36.

Snow, Katherine, and Milisen, Robert. "Spontaneous Improvement in Articulation as Related to Differential Responses to Oral and Picture Articulation Tests." Journal of Speech and Hearing Disorders, (Monograph Supplement 4, 1954b), 46-9.

Sommers, Ronald K.; Cockerille, Clara E.; Paul, Clara D.; Bowser, Colores C.; Fichter, George R.; Fenton, Ann K.; and Copetas, Florence G. "Effects of Speech Therapy and Speech Improvement Upon Articulation and Reading." Journal of Speech and Hearing Disorders, 26 (No. 1, 1961), 27-38.

Spriestersbach, Duane C., and Curtis, James F. "Misarticulation and Discrimination of Speech Sounds." Quarterly Journal of Speech, 37 (No. 4, 1951), 483-91.

Steer, M. D., and Drexler, Hazel G. "Predicting Later Articulation Ability from Kindergarten Tests." Journal of Speech and Hearing Disorders, 25 (No. 4, 1960), 391-7.

Templin, Mildred C. "A Study of Sound Discrimination Ability of Elementary School Pupils." Journal of Speech and Hearing Disorders, 8 (No. 2, 1943), 127-32.

Templin, Mildred C. "Spontaneous Versus Imitated Verbalizations in Testing Articulation in Preschool Children." Journal of Speech and Hearing Disorders, 12 (No. 3, 1947a), 293-300.

Templin, Mildred C. "A Non-Diagnostic Articulation Test." Journal of Speech Disorders, 12 (No. 4, 1947b), 392-6.

Travis, Lee Edward, and Rasmus, Bessie. "The Speech Sound Discrimination Ability of Cases With Functional Disorders of Articulation." Quarterly Journal of Speech, 17 (No. 2, 1931), 217-26.

Weiner, Paul S. "Auditory Discrimination and Articulation." Journal of Speech and Hearing Disorders, 32 (No. 1, 1967), 19-28.

Wilson, Betty Ann. "The Development And Evaluation of a Speech Improvement Program for Kindergarten Children." Journal of Speech and Hearing Disorders, 19 (No. 1, 1954), 4-13.

## Unpublished Materials

Arkebauer, Herbert J., and Ohlman, Patricia. "A Comparative Study of Two Predictive Tests of Articulation." Paper Presented at the 1972 ASHA Convention, San Francisco, California, Nov. 19, 1972.

Chase, Bradley R. "A Comparison of PSTA Scores Assigned by Experienced and Inexperienced Judges." Unpublished study, Eastern Illinois University, 1973.

Hall, William Francis. "A Study of the Articulatory Skills of Children from Three to Six Years of Age." A Dissertation Presented to the Faculty of the Graduate School, The University of Missouri, August, 1962, Copyright, 1963.

Healey, William Charles. "A Study of the Articulatory Skills of Children from Six to Nine Years of Age." A Dissertation Presented to the Faculty of the Graduate School, The University of Missouri, August, 1963, Copyright, 1964.

## Other Sources

Templin, Mildred C. and Darley, Frederic L. The Templin-Darley Tests of Articulation. 2nd ed. Iowa City: State University of Iowa, Bureau of Educational Research and Service, 1969.

Williams, Frederick, ed. Analysis of Production Errors in the Phonetic Performance of School-Age Standard English-Speaking Children; Final Report. Austin: University of Texas, 1971.

Van Riper, Charles. A Predictive Screening Test for Children with Articulatory Speech Defects. Final Report, Cooperative Research Project No. OE-1538, Office of Education, U. S. Department of Health, Education, and Welfare, 1966.

Van Riper, Charles, and Erickson, Robert. Cross-Validation of a Predictive Screening Test for Children with Articulatory Speech Defects; Final Report. U. S. Department of Health, Education, and Welfare: Office of Education, Bureau of Research, 1968a.

Van Riper, Charles, and Erickson, Robert. Predictive Screening Test of Articulation: Manual. Kalamazoo, Western Michigan University, 1968b.