A NUMERICAL
MEASURE OF ARTICULATION

by

Janet O'Neill Barker

A Thesis Submitted to the Faculty of the

DEPARTMENT OF SPEECH

In Partial Fulfillment of the Requirements
For the Degree of

MASTER OF ARTS

In the Graduate College

UNIVERSITY OF ARIZONA

1959

STATEMENT BY AUTHOR

This thesis has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under the rules of the Library.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in their judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Janet O'Neill Barker

APPROVAL BY THESIS DIRECTOR

This thesis has been approved on the date shown below:

James D. Lambert
JAMES D. LAMBERT
Assistant Professor of Speech

18 May 1959
Date

ACKNOWLEDGMENTS

The writer wishes to thank Dr. James D. Lambert for his assistance in the planning and preparation of the thesis. The writer also wishes to thank Dr. Klonda Lynn and Mr. Kenneth Dimmick for advice given during the preparation of the thesis.

Thanks are also extended to the many people who assisted in getting subjects for the study: Mr. Frank Gunther, Director of Special Education at L. M. Prince Elementary School; Mrs. Berta Shires, Principal of Helen Keeling Elementary School; Mrs. Laura Ganoung, Director of Special Education, Tucson Public Schools; Mrs. Connie Green and Mr. Richard Barrett, Tucson Public Schools Speech Therapists; and Keeling and Prince teachers, Mrs. Pagett, Mrs. Hernandez, Miss Runyon, Mr. Brisco, Mr. Smith, Miss Boone, Miss Ackerman, Mrs. Fox, Miss Swenson, and Mrs. Lemon.

Acknowledgment is due Dr. Curtis B. Merritt, College of Education, for statistical advice in the study.

The judges who participated in the study also deserve thanks: Jack Bess, Madellyn Buntz, Jo Clark, Mary Dunham, Diane Kern, Glenda Parrott, Kathleen Shogren, Barbara Votaw, Skip Wallach, and Don Warburton.

The writer also wishes to thank her husband, Boone, for his assistance in editing the tapes and for assistance with the mathematics required.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

BACKGROUND OF THE STUDY

Speech is one of the most important tools in a person's life. It is his main contact with the rest of the world. A child with a speech defect is not only physically handicapped, but psychologically handicapped as well.

Children with speech disabilities constitute one of the largest physically handicapped groups in the world today. The American Speech and Hearing Association Committee on the Midcentury White House Conference on Children and Youth estimated that there were at that time in the United States 2,000,000 seriously speech-defective children between the ages of 2 and 21, or 5 per cent of the assumed population within this age range. Another 5 per cent, according to this survey, had minor speech defects.[1] Various surveys have shown that this may be a low estimate. It is probably safe to say that at least 10 per cent of the elementary- and secondary-school children are suffering

---

[1] ASHA Committee on the Midcentury White House Conference, "Speech Disorders and Speech Correction," Journal of Speech and Hearing Disorders, XVII (June, 1952), pp. 129-130.

from sufficiently serious speech maladies to be considered speech defective.[2]

The various types of articulatory disorders, including those of delayed speech, account for about 75 per cent of the total speech defects found among the public school children.[3] This means that the majority of any correctionist's cases will probably have articulatory problems. Because there are so many children with defects of articulation, it seems desirable to have some easily interpreted means of relating children's articulation to their communication, that is, to the amount of speech listeners can understand. Methods are available for measuring and assigning scores to such intangible factors as intelligence, personality, and adjustment; there should also be a way to assign a single, meaningful score to a child's articulation.

REVIEW OF THE LITERATURE

Any measure of articulation, whether it is to be used for clinical or research purposes, should meet certain criteria:

---

[2] Virgil A. Anderson, Improving the Child's Speech (New York: Oxford University Press, 1953), p. 6.

C. Van Riper, Speech Correction: Principles and Methods, second edition (New York: Prentice-Hall, Inc., 1947), p. 11.

[3] Anderson, Op. Cit., p. 126.

1. it should include a consideration of all speech sounds--consonants, vowels, and diphthongs;
2. it should represent speech adequacy in a quantitative manner;
3. it should be numerically accurate and allow for statistical manipulation;
4. it should be simple and convenient to use;
5. it should be easily interpreted.

A review of the literature of the field of speech correction indicates that people have tried to measure articulation in many ways. The principal method has been to use groups of observers to evaluate subjectively recorded segments of children's speech. Some individual methods have also been used.

Measurement of Speech

Many people have tried to quantify speech for therapeutic and research pruposes. This has usually involved subjective evaluations by observers.

Measurement of Speech by Observation. Most speech therapists measure articulatory defects merely by giving articulation tests and noting the errors. If the number of pre-therapy errors has been reduced, or entirely eliminated, as shown by the post-therapy results, improvement is evident.

This method of measuring defectiveness of articulation is insufficient for research purposes. In cases of multiple articulatory defects it is difficult to compare adequately the speech of different children. Where comparisons are not necessary and where there is no need to quantify a child's improvement, an articulation test before and after therapy is sufficient. When a sound is eliminated from the test record, improvement can be seen. In such cases the nature of the improvement, the intelligibility increase, is not important.

Norma Schneiderman investigated the relationship between language ability and articulation ability.[4] In her study she obtained an articulation score for each child by counting the number of defective sounds which appeared when she used the Test Cards accompanying the Bryngelson and Glaspey Speech Improvement Cards. Although the results of Schneiderman's study are not pertinent to the present investigation, the study is cited to show the common method of evaluating defectiveness of articulation.

Merlin J. Mecham of Brigham Young University constructed a scale for measuring the level of verbal language development

---

[4] Norma Schneiderman, "A Study of the Relationship between Articulation Ability and Language Ability," Journal of Speech and Hearing Disorders, XX (December, 1955), 359-363.

in children.[5] He includes, not only articulation, but all forms of verbal communication behavior: speech, listening, reading, and writing.

Sixty speech items are included in his scale, ranging from the smiles of the very young infant to the ability of the teen-ager to discuss current events with others. The scale is divided by age levels, 0 to 1 year to the 12 to 16 year level. The behavior of the subject must be observed, and his abilities checked on the test form. A plus (+) is given for all items passed, and a minus (-) for every item failed. Each plus rates one point. The total points in that age level having the highest number of consecutive plus points is the basal score. Any points beyond the first minus in this age level are added to the basal score. The total points which each subject receives indicates his approximate age level in verbal communication development.

While this appears to be an excellent scale for measuring the total verbal communication behavior of children, it is not applicable to the measurement of articulation defectiveness. Articulation is not the only speech area included, and cannot be separated from speech as a whole.

---

[5] Merlin J. Mecham, A Scale for Measuring Level of Verbal Communication Behavior in Children (Unpublished Work, Brigham Young University, Utah, 1958).

Measurement of Speech by Comparison. Comparison of speech samples is another subjective method of speech measurement. Comparison can be made of the speech of two different subjects or of two different samples of the speech of the same subject.

Curry, Kennedy, Wagner, and Wilke devised a "product scale" to be used as a basis of comparison for assigning values to articulation.[6] Recordings were made while each of eight children read a selection of nonsense prose. Their articulation ranged from mildly to severely defective. Twenty-five adult observers attempted to rate the speech of these children along a severity continuum. They compared each sample of speech to every other sample in pairs and selected the better of each pair. The first try was unsuccessful because of wide gaps left in the scale. The group as a whole was unable to agree on the relative severity of some of the samples.

In hopes of obtaining better results, a second experiment was designed using ten recorded speech samples and 139 observers. Neither the subjects nor the observers were the same as in the previous experiment. Again, lack of agreement left wide gaps in the scale.

---

[6] Robert Curry, et al., "A Phonographic Scale for Measurement of Defective Articulation," Journal of Speech Disorders, VIII (March, 1943), 123-126.

Rather than conduct another experiment, a third group of fifty-three observers was asked to judge those pairs upon which no final decision had been reached by the second group of observers. Sufficient agreement was obtained in this manner to complete the "product scale."

The scale can be used, according to the writers, by comparing the speech of a person reading the same selection of nonsense prose with the recorded samples of speech. The scale value of the selection most resembling the person's speech is then assigned to him.

Although the "product scale" designed by Curry and others may be a valid measure of articulation defectiveness, it is felt that more thorough investigation is necessary before it can be accepted. Only eight children were selected for the first experiment, and ten were used for the second. This is a small sample to be used to construct such a scale.

Measurement of Speech by Judgment Ratings. A method of measuring speech improvement was devised by George Shames in his investigation of the value of various biographical and personality factors as predictors of success in speech therapy.[7] He studied a group made up of twenty-seven stutterers, two subjects with

---

[7] George Shames, "An Investigation of Prognosis and Evaluation in Speech Therapy," Journal of Speech and Hearing Disorders, XVII (December, 1952), 386-392.

voice problems, and eight with articulatory defects. Four of this last group were cleft palate cases.

Problems of people with defective speech appear to manifest themselves through speech inadequacy and social inadequacy; therefore, the success of speech therapy was evaluated in terms of the alleviation of symptoms of these factors. Speech adequacy was measured by presenting randomly pre- and post-therapy recordings of speech to two speech clinicians who were unfamiliar with the subjects. They were asked to indicate whether the speech on the second recording was better than, the same as, or worse than that on the first recording. The correlation between the judgments of the two clinicians was .57.

The present study is not concerned with the final results of Shames' work. Only a consideration of the methods used to measure speech adequacy is applicable. The fact that the group tested was unbalanced as far as the distribution of types of speech problems was concerned tends to make the value of the entire study questionable. Furthermore, the correlation of the two judges' ratings was not very high (.57). It appears that judgments of two observers are not sufficiently reliable in determining articulation defectiveness.

Several studies have been done at Ohio State University using judges' ratings of recorded speech. These judges were trained according to the methods determined by Sheila Morrison.[8] Using the method of equal appearing intervals, Morrison constructed a severity scale of articulation defectiveness. There were two parts to her study--a preliminary investigation and a major investigation.

In her preliminary investigation forty naive and twelve experienced observers were asked to place tape recorded segments of the speech of sixty children along a severity continuum, ranging from one, for least severe, to nine, for most severe articulation defectiveness. The children were five to ten years of age. Five-, ten-, and fifteen-second speech segments were used. No evidence was presented to indicate that any of the segment lengths produced more reliable results than either of the other two. The purpose of the preliminary investigation was to determine the best length of a segment of speech which could be reliably placed along a severity continuum.

Since the experts expressed a preference for either the five- or ten-second segments, these were used for the major investigation--the actual construction of the severity scale. Two groups of observers, naive and expert, were again used. There were thirty-two to thirty-six observers in each group. The observers were again asked to place the samples of speech along

---

[8] Sheila Morrison, "Measuring the Severity of Articulation Defectiveness," Journal of Speech and Hearing Disorders, XX (December, 1955), 347-351.

a <u>one</u> to <u>nine</u> severity continuum. There was high agreement between the two groups with respect to placement of the speech segments along the severity scale, although the experts tended to rate the segments more severely than did the naive observers.

In a subsequent study, Sheila Morrison and Dorothy Sherman used the original Morrison samples of speech to determine the reliability of individual ratings of articulation defectiveness.[9] One-minute samples of the speech of fifty of Morrison's sixty children with varying degrees of articulation defectiveness were judged by two groups of ten observers. The samples were presented to one group in five-second segments, to the other in ten-second segments. A nine-point scale was used, with <u>one</u> defined as least severe articulation defectiveness and <u>nine</u> as most severe. Each group was trained with Morrison's severity scale.

The writers concluded that speech samples can be reliably placed in relative positions along a severity continuum by a trained individual observer, even though absolute ratings of the severity of defective articulation are not necessarily comparable from one individual observer to another. Observers differed

---

[9] Dorothy Sherman and Sheila Morrison, "Reliability of Individual Ratings of Severity of Defective Articulation," <u>Journal of Speech and Hearing Disorders</u>, XX (December, 1955), 352-358.

significantly in their general level of rating. Sherman and Morrison also concluded that results would be equally reliable when either five- or ten-second speech segments were used.

Another investigation was conducted at Ohio State University to determine the reliability of the Morrison rating scale.[10] The split-half product-moment correlation coefficient was used with ten, eight, six, four, and two judges. The Sherman-Brown prophecy formula was then applied to the observed values for predicted correlations of five, four, three, two, and one judges. The results showed a correlation of .95 for five judges, .94 for four, .90 for three, .85 for two, and .58 for one. The Edgerton-Troops average correlation did not yield any more significant information than was obtained by the Pearson Product-Moment Correlation Coefficient. Analysis of variance according to the Ebel and Hoyt revision of the Horst formula also produced similar results.

The split-half and analysis of variance procedures were used to determine the reliability of the test items. The combined ratings of ten judges to odd-numbered items were correlated with the combined responses of ten judges to even-numbered items. This resulted in a reliability of .95. The same

---

[10] Nancy March, et al., "Observed and Predicted Estimates of Reliability of a Speech Articulation Rating Scale," <u>Speech Monographs</u>, XXV (November, 1958), 296-403.

prediction formula used to determine the number of judges necessary to achieve a reliability rating of .90 was applied to the number of test items required to achieve this reliability, and yielded 2.0 items.

Attempts to determine the reliability of the number of subjects were not entirely successful. The split-half method appeared to tell more about the judges than about the number of subjects. However, analysis of variance did offer some possibility for determining the reliability of the number of subjects. In general, an upward trend was shown when the number of subjects was increased from five to forty in groups of five. A reliability of .97 was obtained for forty subjects.

The investigators concluded that a reliability of .90 would be obtained if either ten judges and two items for each subject or six judges and four items for each subject were used.

Measurement of Speech by a Numerical Score. There has been one attempt to assign a numerical score of defective articulation. This was made by Kenneth Scott Wood in his study of parental maladjustment and functional articulatory disorders.[11] Wood assumed that since some sounds appear very seldom in the English language, inability to produce such sounds cannot be

---

[11] Kenneth Scott Wood, "Parental Maladjustment and Functional Articulation Defects in Children," Journal of Speech Disorders, XI (December, 1946), 149-153.

considered as severe a problem as inability to produce more frequently used sounds. He asserted that consonant sounds played the most important role in speech intelligibility. He, therefore, constructed an Articulation Index based on Travis' table of the frequency of occurrence of consonant sounds in the speech of American Children.[12] Travis' study will be discussed more fully in a later section (see Table I). Wood prorated the sound values of Travis' table into the initial, medial, and final positions in words. Each of Wood's subjects was given an articulation test, and all sound errors were noted. The numerical values of the sounds correctly produced were then added together to obtain a single score. The score that each child received, the Articulation Index, was a quantitative description of his ability to articulate sounds correctly.

Ernest Henrickson criticized Wood's equal prorating, showing that consonant sounds did not appear an equal number of times in all positions.[13] His findings concerning the frequency of occurrence of consonant sounds did, however, approximate those of Travis.

---

[12] Lee Edward Travis, Speech Pathology (New York: D. Appleton and Company, 1931), p. 223.

[13] Ernest H. Henrickson, "An Analysis of Wood's Articulation Index," Journal of Speech and Hearing Disorders, XIII (September, 1948), 233-235.

In defense of his method, Wood pointed out that the positions of consonant sounds were not stable in continuous speech.[14] The position of a sound is relative, not only to its position in a word, but to the placing of the word in a sentence, and the phrasing of the sentence. For example, a sound which is normally final, as the /t/ in hat, is medial in the sentence, "My new hat is red;" whereas, it is final in the sentence, "I just bought a new hat, which I like very much." Wood also stated that prorating the sounds into positions allowed partial credit for partial learning in speech therapy. He further suggested that a person partially learns a sound in all positions when he learns it in one, even though he may not be able to produce that sound correctly in the other positions at a given time.

Upon first examination the studies at Ohio State University[15] appeared to meet the need for a measure of articulation. They showed that trained judges could reliably place speech samples along a one to nine severity continuum. They

---

[14] Kenneth Scott Wood, "Measurement of Progress in the Correction of Articulatory Speech Defects," Journal of Speech and Hearing Disorders, XIV (June, 1949), 171-174.

[15] Morrison, Op. Cit.

Sherman and Morrison, Op. Cit.

March, et al., Op. Cit.

found that ten judges rating two samples of speech or six judges rating four samples of speech yielded a reliability coefficient of correlation of .90. Judges' ratings of defective articulation recognize the relation of articulation to speech adequacy and include all speech sounds.

Using judges to rate samples of speech is impractical as a clinical and research tool because of the large number of observers and speech samples necessary to achieve reliable ratings. Furthermore, it is difficult to get adequate samples of children's speech. Background noise, poor acoustics, or technical problems reduce the value of tape recorded speech samples. If the material to be recorded is to be read, it must be selected to fit the age groups being tested. The material must be simple enough for the poorest reader and interesting enough for the most mature reader. Sometimes several samples of written material are necessary to fit all levels of readers and to eliminate memorizing on the part of the judges. The last subject has an advantage over the first if every subject reads the same selection. The judges may unconsciously fill in all omitted sounds or become so bored that they will not listen carefully to the last samples presented. If speech is used, the topics discussed must be fairly standard. If the judges cannot tell what the subject is talking about, he may receive a poorer rating than he deserves. A ten-second sample of speech is not very long to be taken from the middle of a conversation. Sometimes the subjects have difficulty thinking of something to say. It is sometimes necessary to spend fifteen or twenty minutes in order to get a minute of connected speech.

Other problems arise when it is time for the judges to meet. The more observers required, the more difficult it is to get them all together at once. Judges must meet in a room where a tape recorder can be heard by all of them. The room has to be quiet and free from distractions. A five- or ten-second sample of speech is sufficient to obtain a reliable estimate of the articulation defectiveness present on the sample.[16] However, if distractions are present or if the samples cannot be heard, they may have to be played several times before all observers feel their ratings are accurate.

Many hours are required to train the judges and to tabulate final results. Such a time consuming procedure, even though reliable, will not encourage research in the area of articulatory defects.

Ratings by judges are not absolute; they are made on a relative basis. For example, a rating of nine given to two different children does not necessarily mean the same thing in each case. If an entire group of subjects is compared at one time, the rating of nine is given to the child whose articulation is most severely defective in the group. If a group is not being rated all at once, the observers must rely on experience, considering

---

[16] Morrison, *Op. Cit.*

nine as the worst articulation each has heard. One other possibility is to have a standard tape showing all possible levels of severity of defective articulation to be used as a basis for comparison. In a clinical situation rating all cases every time a report was required on one would be highly impractical. Unless some type of standard tape were provided, ratings would be meaningful to no one except the observer.

A numerical score does not have the drawbacks of judgment ratings. The subjective evaluation is reduced to a minimum. It is easier to know whether a child misarticulates or omits a sound than to try to decide where he fits on a severity continuum. The order of the subjects does not affect the score because a subjective evaluation of placement is not made. A moderate amount of noise and a few distractions do not markedly hinder the progress of an articulation test. The tester can compute the scores for each subject on his own time and does not need to meet the schedules of several judges. If a numerical score is based on the frequency of occurrence of speech sounds, it can also be said to relate to speech adequacy. A numerical score is the only method of measuring articulation which can be statistically manipulated.

Wood's Articulation Index, the only numerical method of measurement which has been devised, includes consonants only.[17]

---

[17] Wood, "Measurement of Progress in the Correction of Articulatory Speech Defects," Op. Cit.

If a score is to represent accurately speech adequacy, it must include all speech sounds. For this reason the frequency of occurrence studies will be reviewed in an effort to find a complete sound distribution table.

Frequency of Occurrence Studies

By counting the frequency of occurrence of the English phonemes, various studies have given each sound a numerical value based on a final total of approximately 100. The frequency of occurrence tables show the probability that a given phoneme will occur in a segment of speech or written material one hundred sounds in length.

Frequency of Occurrence of Sounds in Written Material. The earliest study of the frequency of occurrence of English sounds was reported by Godfrey Dewey in 1923.[18] Dewey tabulated consonants, vowels, and diphthongs found in such written material as advertising, speeches, current periodicals, and personal and business correspondence (see Tables II and III). The results of his study later became the foundation for the PB (Phonetically Balanced) word lists which are commonly used in auditory speech reception tests.

---

[18] Godfrey Dewey, Relativ Frequency of English Speech Sounds (London: Humphrey Milford, Oxford University Press), Cambridge: Harvard University Press, 1931 , p. 123.

Frequency of Occurrence of Speech Sounds. All other studies of the frequency of occurrence of English or American sounds have been based on speech. Studies have included both children and adults.

Lee Edward Travis made an extensive study of the English language and compiled tables of the frequency of occurrence of consonant sounds in the speech of children, adult laborers, and university adults.[19] Travis' results are shown in Table I. The "original" and "Corrected" columns for the speech of university adults figures need to be explained. Travis reported that the total number of sounds counted for the speech of university adults was 16,077;[20] his percentages were compiled from this figure. However, it was found that the total should have been 16,226. All percentages were re-calculated from the given data.[21]

Charles H. Voelker counted the frequency of occurrence of vowels and consonants in radio speech.[22] A total of 5,946 radio announcements, of at least one hundred sounds in length,

---

[19] Travis, Op. Cit., pp. 223-225.

[20] Travis, Op. Cit., p. 224.

[21] Ibid.

[22] Charles H. Voelker, "Technique for a Phonetic Frequency Count in Formal American Speech," Extrait des Archives Neerlandaises de Phonetique Experimentale, XI (1935), 69-72.

TABLE I

TRAVIS FREQUENCY OF OCCURRENCE TABLES

Sound	Children	Laborers	University Adults	
			original	corrected
/t/	12.00	12.50	13.40	13.24
/n/	10.40	10.80	10.30	10.14
/r/	9.30	8.80	8.80	8.70
/s/	8.90	7.20	7.40	7.36
/d/	6.30	6.80	7.00	6.91
/l/	6.30	5.50	6.00	5.91
/m/	5.20	4.70	4.90	4.89
/k/	5.10	5.40	4.80	4.73
/ð/	4.00	4.50	4.30	4.27
/w/	4.20	4.60	4.30	4.24
/h/	3.90	4.30	4.20	4.22
/z/	4.30	3.60	4.00	4.05
/f/	2.40	2.70	2.70	2.71
/b/	2.90	2.80	2.60	2.64
/v/	2.40	2.60	2.60	2.57
/g/	2.70	2.50	2.50	2.52
/j/	1.70	2.10	2.40	2.42
/p/	2.80	2.00	2.30	2.24
/ŋ/	1.90	1.90	1.90	1.88
/ʃ/	1.13	1.40	1.60	1.56
/θ/	0.90	0.70	1.20	1.10
/dʒ/	0.70	0.60	0.70	0.64
/tʃ/	----	0.90	0.60	0.56
/hw/	0.60	0.30	0.50	0.51
/ʒ/	0.06	0.08	0.02	0.02
Total	100.79	99.28	101.02	100.03

was recorded. These announcements included reading, oratory, and extemporaneous and impromptu speech. Both prose and poetry were represented. The same announcer was recorded no more than three times. Whenever an announcer was recorded more than once, he presented a different type of material than he had presented on the previous sample. Most of the announcers had college training, theatrical experience, or training in speech, music, or dramatics. They came from all sections of the United States, and some of them were from homes where a language other than English was predominantly spoken. A table of vowels and consonants was compiled showing the percentage of times each sound occurred within the 660,594 sounds recorded (see Table II).

The most complete study of the relative frequency of occurrence of sounds in American speech was conducted at Bell Laboratories by Norman French, Charles Carter, Jr., and Walter Koenig, Jr.[23] The material for the study was obtained from a series of telephone conversations over typical toll circuits which terminated in the city of New York. The words used in the conversations were noted in the following manner: only the nouns were recorded one week; the next week the observer recorded all verbs; only adjectives and adverbs were recorded during the third

---

[23] Norman R. French, Charles W. Carter, Jr., and Walter Koenig, Jr., "The Words and Sounds of Telephone Conversations," Bell System Technical Journal, IX (April, 1930), 290-324.

TABLE II

FREQUENCY OF OCCURRENCE OF SOUNDS

Sound	Voelker	Dewey (converted)	Bell Laboratories (½ of total)
/t/	11.66	11.48	14.42
/n/	11.85	11.65	8.75
/r/	10.51	11.09	7.91
/s/	7.54	7.33	5.61
/d/	8.28	6.95	7.16
/l/	6.32	6.02	6.35
/m/	4.47	4.48	5.18
/k/	4.15	4.36	4.81
/ð/	5.13	5.52	3.98
/w/	2.99	3.35	4.69
/h/	2.66	2.91	2.87
/z/	3.48	4.79	3.46
/f/	3.48	2.96	2.97
/b/	3.18	2.91	2.64
/v/	2.52	3.67	2.75
/g/	1.75	1.19	2.35
/j/	1.89	0.97	3.24
/p/	2.41	3.28	2.60
/ŋ/	1.68	1.54	1.78
/ʃ/	1.64	1.32	1.03
/θ/	1.06	0.60	1.03
/dʒ/	----	0.71	0.48
/tʃ/	----	0.84	0.54
/hw/	0.60	----	0.52
/ʒ/	0.67	0.08	0.01
others	----	----	2.37
	99.92	100.00	99.50

TABLE II (continued)

Sound	Voelker	Dewey (converted)	Bell Laboratories
/ə/	17.76	12.20	16.30
/I/	20.56	22.40	15.05
/ɛ/	7.98	9.05	7.69
/aI/	----	4.18	7.58
/ae/	8.06	10.40	6.89
/ɑ/ /ɒ/	10.85	8.68	6.52
/i/	7.40	5.58	6.44
/u/	5.21	4.21	6.26
/eI/ /e/	5.21	4.84	4.78
/oU/ /o/	5.54	4.79	4.74
/ɚ/	----	----	4.56
/ɔ/	5.58	3.32	4.15
/ʌ/	1.12	6.14	4.14
/U/	4.60	1.82	2.96
/aU/	----	1.55	1.69
/ɝ/	----	----	0.80
/ju/	----	0.82	0.26
/ɔI/	----	0.24	0.19
	99.87	100.22	100.00

week. This procedure was repeated until observations had been made for five hundred conversations for each of the three groups of words--nouns, verbs, and adjectives and adverbs. Pronouns, articles, and prepositions and conjunctions were noted in additional conversations. However, it was decided that 150 conversations in each case was sufficient. Certain classes of words--proper names, titles, interjections, letters, numbers, and vocalized pauses such as er or uh--were omitted entirely from the frequency count.

Words recorded by this sampling technique were not exactly the same as the words which would have been obtained had the entire conversation been recorded in each case. However a different observer later recorded all verbs in 250 conversations as a check against the already recorded verbs. The word lists compiled by the two different observers corresponded closely. A table was compiled containing 737 words which were heard in at least 1 per cent of the conversations; this list formed 96 per cent of the total occurrences of all words. It was also found that only 2,240 different words were used in the conversations, which consisted of 80,000 words. The omission of the 1,500 least common words changed the sound distribution by negligible amounts.

All forms of the same word--plural forms of nouns, different tenses of verbs, and comparative and superlative forms of adjectives--were counted as the same word. These forms were recorded, however, and were treated individually when the speech sound analysis was made.

The syllable was used as the basis for the sound frequency analysis. Words were divided into phonetic syllables, not orthographic syllables. The dictionary was not used as the authority for either syllabification or pronunciation. Such words as every, preference, and average were divided into two syllables according to colloquial pronunciation. The unstressed pronunciations of such words as and, to, and of were used.

A single pronunciation was assigned to each of the other recorded words according to the most typical pronunciation in reasonably enunciated speech among the educated people of New York. For instance, the vowel sound in pot and palm was considered the same, even though there is a real distinction in New England pronunciation. Path, can't, last, and ask were classified with pan. The vowel sounds in pair and par were considered to be separate classes.(see Tables II, III, and IV).

The different values in Tables II and IV for Dewey's study and the sound frequency analysis made by French, Carter, and Koenig at Bell Laboratories need to be explained. The Bell Laboratories figures, as they originally appeared (Table III), showed the vowels, initial consonants, and final consonants in three separate categories. They were tabulated on the basis of 100 per cent for each. The vowels and consonants in Dewey's study totaled 100 per cent, the consonants constituting

TABLE III

RELATIVE OCCURRENCE OF SPEECH SOUNDS

ACCORDING TO FRENCH, CARTER, AND KOENIG

All Words Except Articles

Consonants

Sound	Percentage Initial	Percentage Final	Sound	Percentage Initial	Percentage Final
/t/	7.86	14.30	/hw/	0.91	----
/n/	4.99	12.52	/dʒ/	0.83	0.14
/r/	2.78	13.05	/ʒ/	0.02	0.01
/d/	6.21	4.44	tr	0.69	----
/l/	4.31	8.40	nt	----	4.40
/m/	5.89	5.48	ts	----	1.11
/s/	5.46	3.13	kt	----	0.42
/k/	5.55	2.85	st	0.87	1.18
/w/	9.38	----	sp	0.19	----
/ð/	6.72	1.25	ks	----	0.47
/z/	0.34	6.01	nd	----	2.56
/f/	3.96	1.37	ld	----	0.75
/j/	6.48	----	rd	----	0.37
/h/	5.75	----	kw	0.28	----
/v/	1.25	4.23	nk	----	0.76
/b/	4.64	0.42	kl	0.18	----
/p/	2.54	1.24	fr	0.62	----
/g/	4.33	0.38	bl	0.23	----
/ŋ/	----	3.57	pr	1.06	----
/ʃ/	1.74	0.32	pl	0.36	----
/θ/	2.02	0.04	rz	----	0.57
/tʃ/	0.55	0.53	others	1.01	3.73
	92.75	83.53		7.25	16.47

200.00

Total number of sounds: Initial: 64,043 Final: 65,544

TABLE III (continued)

Vowels			
Sound	Percentage	Sound	Percentage
/ə/	5.52	/i /	6.44
	5.33	/u /	6.26
	2.65	/eI/ /e /	4.78
	1.83	/oU/ /o /	4.74
	0.97	/ɚ/	4.56
/I /	10.27	/ɔ /	4.15
	3.78	/ʌ/	4.14
/ɛ/	6.60	/U /	2.96
/ɛɚ/	1.09	/aU/	1.69
/aI/	7.58	/ɝ /	0.80
/ae/	6.89	/ju/	0.26
/ɑ/	5.21	/ɔI/	0.19
/ɑɚ/	1.31		
		Total	100.00

Total number of sounds:

92,522	Vowels
64,043	Initial Consonants
65,544	Final Consonants
222,109	

TABLE IV

COMPARISON OF DEWEY AND BELL LABORATORIES FIGURES

Consonants			Vowels		
Sound	Dewey	Bell Laboratories (converted)	Sound	Dewey	Bell Laboratories (converted)
/t/	7.13	9.01	/ə/	4.63	6.80
/n/	7.24	5.13	/I/	8.53	5.86
/r/	6.88	4.65	/ɛ/	3.44	3.21
/d/	4.31	4.21	/aI/	1.59	3.16
/l/	3.74	3.72	/ae/	3.95	2.87
/m/	2.78	3.32	/ɑ/ /ɒ/	3.30	2.72
/s/	4.55	3.32	/i/	2.12	2.69
/k/	2.71	3.03	/u/	1.60	2.61
/w/	2.08	2.70	/eI/ /e/	1.84	1.99
/ð/	3.43	2.31	/oU/ /o/	1.63	1.98
/z/	2.97	2.04	/ɚ/	----	1.90
/f/	1.84	2.11	/ɔ/	1.26	1.73
/j/	0.60	1.87	/ʌ/	2.33	1.73
/h/	1.81	1.66	/U/	0.69	1.23
/v/	2.28	1.61	/aU/	0.59	0.70
/b/	1.81	1.52	/ɝ/	----	0.33
/p/	2.04	1.51	/ju/	0.31	0.11
/g/	0.74	1.47	/ɔI/	0.09	0.08
/ŋ/	0.96	1.05			
/ʃ/	0.82	0.59			
/θ/	0.37	0.59			
/tʃ/	0.52	0.32			
/hw/	----	0.26			
/dʒ/	0.44	0.28			
/ʒ/	0.05	0.02			
	62.10	58.30		37.90	41.70

Dewey Total: 100.00 | Bell Laboratories Total: 100.00

62 per cent of the total percentage of sounds. Both sets of figures (Dewey's and Bell Laboratories') were converted so they could be compared on an equal basis to each other and to the other frequency of occurrence studies. Only these two sets of figures were converted, since they alone considered vowels, diphthongs, and consonants.

The Bell Laboratories study divided consonant sounds into the initial and final positions in the syllable. The values of the sounds in each position totaled 100 per cent. In order to compare the Bell Laboratories figures on an equal basis to all other figures (Dewey, Travis, Voelker), which used 100 per cent as the total for all consonant sound values, the total number of times a consonant occurred in both positions was combined (giving a total of 200 per cent) and divided by two. The relative frequency of each consonant sound was the same; the total values for each sound were cut in half.

When the Bell Laboratories figures were converted, the blends were combined with the consonants. The initial consonants in initial blends were added to the initial consonant total; the final consonant in the final blends were added to the total of the final consonants. Thus, the percentage for /st/ in the initial column was added to /s/. In the final column it was added to /t/.

By dividing the number of sounds in each category--vowels, initial consonants, and final consonants--into the total number of sounds recorded in the Bell Laboratories study, the conversion factors of 0.417 for vowels, 0.288 for initial consonants, and 0.295 for final consonants were determined. Multiplication of each sound by its respective conversion factor gave the frequency of occurrence of each sound relative to every other sound. The total of the converted figures equals 100 per cent (see Table IV).

The Dewey figures were converted to compare with the other tables of figures by first dividing 38 per cent (the total vowel percentage) and 62 per cent (the total consonant percentage) into one hundred. The resulting figures, 2.632 and 1.613, were then multiplied by each vowel and consonant value, respectively. The results gave a total of 100 per cent for each column--vowels (and diphthongs) and consonants (see Table II).

The frequency of occurrence tables shown in this chapter indicate that there is little difference in the frequency of occurrence of sounds, whether the study is based on spoken or written material. The most complete study was done by French, Carter, and Koenig at Bell Laboratories. This study included a frequency count of vowels, consonants, and diphthongs.

CHAPTER II

STATEMENT OF THE PROBLEM

The present investigation was designed to study the relationship of an Articulation Score based upon the relative frequency of occurrence of all speech sounds and proficiency of articulation as judged by experienced speech therapists. The basic assumption underlying this investigation is that if such an Articulation Score correlates highly with competant subjective evaluations of an individual's articulatory proficiency, then the Articulation Score is a valid measure of articulatory proficiency. Specifically, the study sought to answer the following questions:

1. Can a scale be devised which will yield a single score which is numerically accurate, statistically manipulative, easily understood, and convenient and simple to use?
2. Is such a score valid as a representation of articulatory proficiency?

IMPORTANCE OF THE STUDY

Since so many of the speech difficulties of children are articulatory in nature, research in this area is needed. With the exception of the investigations of the causes of articulatory disorders, little research has been done. The primary reason is that no adequate method of measuring articulation is available. A method needs to be devised which can be statistically manipulated and which is meaningful, accurate, convenient, and relatively simple to use.

Such a measure also has implications for the speech therapist. A convenient quantitative measure of the child's articulation is often desirable. Parents, teachers, doctors, or psychologists may request periodic reports of the child's progress. If a simple defect is present, such as the substitution of /θ/ for /s/, it is not difficult to indicate that the child originally substituted /θ/ for /s/ in all positions and that at a given time he can produce the /s/ sound in the medial and final positions, misarticulating only the initial sound. However, in multiple articulatory defects, such as are found in some children with Cerebral Palsy, such a method of analysis is long, clumsy, and inefficient. A total picture of the child's speech problem is also difficult to present with explanation and description only. Counting the phonemes which a person produces correctly in continued speech gives a rough numerical evaluation of articulatory ability. It would be better to have a single numerical score, representing the child's total articulatory proficiency. If it can be said that a child's articulation was 60 per cent proficient at the beginning of therapy and was 90 per cent proficient after a given period of time, a more comprehensive and more easily understood measure is presented. Such a score may be supplemented with explanatory material if desired; however, the Articulation Score by itself is a more efficient indication of the child's total speech picture than a longer, more detailed description of each individual sound in each position.

DEFINITION OF TERMS

Speech Defect. Wendell Johnson has provided the simplest and most practical definition of a speech defect: "a child's speech is defective when most listerners pay as much attention, or more, to how he speaks as to what he says."[24]

Articulation. Articulation concerns the distortion, substitution, or omission of sounds. Both consonant and vowel sounds should be included in any consideration of articulation.

Articulation Defectiveness. The amount of the child's speech, i.e., the sounds that come out of his mouth, which is not easily understood because of faulty articulation is known as articulation defectiveness. Articulation defectiveness is related to communication. The more defective a person's articulation is, the more difficulty others will have understanding him, and the less information will be received from his speech.

---

[24] Wendell Johnson, et al., Speech Handicapped School Children, (New York: Harper and Brothers Publishers, 1948), p. 2.

Articulation Score. The Articulation Score is a single numerical score given to an individual's articulation on the basis of the sounds misarticulated or omitted on an articulation test. The score indicates the total proficiency of a person's articulation.

Reliability. Reliability is a function of the internal consistency of the measuring instrument. There are three methods of determining reliability: the split-half method, the alternate or parallel forms method, and the test-retest (repetition) method.[25]

Validity. Validity is concerned with the question, "Does this measure what it is supposed to measure?" Validity is determined by correlating the method of measurement in question with a measure of known or assumed validity.[26]

---

[25] Henry E. Garrett, Statistics in Psychology and Education, fourth edition (New York: Longmans Green and Co., 1957), p. 332.

[26] Garrett, Op. Cit., pp. 344-345.

CHAPTER III

PROCEDURE

The purpose of this study was to construct an Articulation Score representing a person's articulatory proficiency. This score must fulfill all of the criteria mentioned in Chapter I. The problem was resolved through the following procedures:

1. The literature pertinent to the study was reviewed to discover the existing methods of measuring articulation.
2. An evaluation of these methods was made.
3. Studies concerning the frequency of occurrence of sounds were reviewed, and the one which was the most complete was selected as the basis for the Articulation Score.
4. An articulation scale was constructed which yielded a single numerical score representative of a person's defective articulation.
5. A group of children with articulatory problems was tested, and an Articulation Score was assigned to each of them.
6. Trained speech therapists evaluated samples of the speech of these children and gave each a rating of articulation defectiveness.

7. The judges' evaluation of each subject's articulation was correlated with his Articulation Score.

The best way to relate a person's articulatory proficiency to the amount of speech his listeners could understand was to assign each speech sound a value based on the number of times that sould would probably be heard in one hundred consecutive American speech sounds. This was the assumption which prompted Wood to base his Articulation Index on Travis' frequency table.[27] Wood included only consonants in his study. Vowel sounds are also misarticulated. In some cases, vowel sounds present the greatest problem. For example, the Spanish speaking person often has great difficulty distinguishing between the American /i/ and /I/ or saying /æ/. This can cause the listener as much confusion as substitution of /t/ for /θ/ or /d/ for /ð/.

The Bell Laboratories study [28] included all of the vowels, diphthongs, and consonants of American speech. An estimate of the accuracy of these figures may be made from the information presented in Table V. These data were compiled from the check test mentioned in Chapter I, conducted on verbs alone. The

---

27 Wood, "Parental Maladjustment and Functional Articulation Defects in Children," Op.. Cit.

Travis, Op. Cit.

28 French, Carter, and Koenig, Op. Cit.

TABLE V

BELL LABORATORIES CHECK TEST

Relative Occurrence of Consonants in Verbs

Sound	First Observations	Check Test	Difference
/b /	1.02	1.02	.00
/d /	4.46	4.83	+.37
/f /	1.73	2.18	+.45
/g /	11.15	9.40	-1.75
/h /	1.40	1.66	+.26
/dʒ/	0.22	0.23	+.01
/k /	8.90	8.74	-.16
/l /	7.70	7.94	+.24
/m /	4.45	3.96	-.49
/n /	6.87	7.10	+.23
/p /	3.34	3.47	+.13
/r /	3.97	3.88	-.09
/s /	10.31	9.67	-.64
/t /	16.97	17.39	+.42
/v /	2.36	2.20	-.16
/w /	4.87	4.54	-.33
/j /	0.55	0.53	-.02
/z /	1.09	1.44	+.35
/ʧ/	0.32	0.75	+.43
/ð/	1.35	1.17	-.18
/θ /	2.53	2.85	+.32
/ʃ/	0.05	0.06	+.01
/ʒ /	0.00	0.00	.00
/ŋ /	4.39	5.00	+.61
	100.00	100.00	

check test was made on the same set of toll circuits, but by a different observer and at a different time. The check test showed a close relationship between the original sound count and the later figures.

ARTICULATION SCORE

Since the Bell Laboratories study was the most complete, and since it appeared to be the most accurate, the frequency of occurrence tables resulting from this study were used as the basis for the Articulation Score in the present investigation.

Wood's reasons for equal prorating of consonants were thought to be valid.[29] To give partial credit for partial learning of a sound, it is necessary to have some division of consonant sounds. Most speech therapists use the three word positions--initial, medial, and final--in articulation tests and therapy. For these reasons the total consonant values of the Bell frequency analysis were prorated into three word positions. Sounds which do not normally appear in three positions were prorated according to the number of positions in which they can be found in American speech. Thus, /t/ was prorated into initial, medial, and final; while /h/ was prorated into two positions, initial and medial.

---

[29] Wood, "Measurement of Progress in the Correction of Articulatory Speech Defects," Op. Cit.

For purposes of scoring in the present study a few of the original Bell Laboratories figures were combined into like classes. The /ɑ/ in par was added to /ɑ/ in pot; /ɛ/ in pair was added to /ɛ/ in pen. In the same manner /I/ in receive was added to /I/ in pin. The unstressed vowel sounds in possible, about, notion, wanted, and people were combined into the single category /ə/.

In the category of the blends a small proportion (1.39 per cent) was assigned to "others." The representativeness of the word list published in the study report has been mentioned previously (cf page 24). Because this list of words represented 96 per cent of all words recorded, it was assumed that the distribution of the 1.39 per cent of unassigned blends would be sufficiently accurate if the blends not previously noted but appearing in the word list were to be counted and proportionately divided into 1.39 per cent. There were 197 occurrences of new blends in the word list. The percentage of times each blend occurred in the list was added to the original figures for the consonant concerned (cf page 29).

Each sound was assigned a value according to its frequency of occurrence in American speech (see Table III). The total values of the misarticulated sounds was the Articulation Score. The Articulation Score, therefore, represented the inadequacy, in arithmetical form, of the subject's articulation.

TABLE VI

SOUND VALUES FOR THE ARTICULATION SCORE

Equal Prorating by Word Position (AS)

Sound		Values: Initial	Medial	Final	Total
/t /		3.00	3.00	3.00	9.00
/n /		1.71	1.71	1.71	5.13
/r /		1.55	1.55	1.55	4.65
/d /		1.41	1.41	1.41	4.23
/l /		1.24	1.24	1.24	3.72
/s /		1.11	1.11	1.11	3.33
/m /		1.10	1.10	1.10	3.30
/k /		1.01	1.01	1.01	3.03
/w /		1.35	1.35	----	2.70
/ð/	THen	0.77	0.77	0.77	2.31
/z /		0.68	0.68	0.68	2.04
/f /		0.70	0.70	0.70	2.10
/j /		0.93	0.93	----	1.86
/h /		0.83	0.83	----	1.66
/v /		0.54	0.54	0.54	1.62
/b /		0.51	0.51	0.51	1.53
/p /		0.50	0.50	0.50	1.50
/g /	Gun	0.49	0.49	0.49	1.47
/ŋ /	NG	----	0.53	0.53	1.06
/ʃ /	SH	0.20	0.20	0.20	0.60
/θ /	THin	0.20	0.20	0.20	0.60
/tʃ/	CH	0.10	0.10	0.10	0.30
/hw/	WHen	0.13	0.13	----	0.26
/dʒ/	JudGe	0.09	0.09	0.09	0.27
/ʒ /	ZH	0.01	0.01	0.01	0.03
		20.16	20.69	17.45	58.30

TABLE VI (continued)

Vowels					
Key Word	Sound	Value	Key Word	Sound	Value
about	/ə/	6.80	pole	/oU/ /o/	1.98
pin	/I/	5.86	differ	/ɚ/	1.90
pen	/ɛ/	3.21	pawn	/ɔ/	1.73
pine	/aI/	3.16	pun	/ʌ/	1.73
pan	/ae/	2.87	pull	/U/	1.23
pot	/ɑ/	2.72	pout	/aU/	0.70
peel	/i/	2.69	purr	/ɝ/	0.33
pool	/u/	2.61	pew	/ju/	0.11
pane	/eI/ /e/	1.99	poise	/ɔI/	0.08
			Total		41.70

Initial Consonants	20.16
Medial Consonants	20.69
Final Consonants	17.45
Vowels	41.70
Total	100.00

For example, if a child substituted /w/ for /r/ in the words ruler, carrot, and car, and /w/ for /l/ in lamb, umbrella, and doll, and omitted /s/ in Santa Claus, he would have missed /r/ and /l/ in all positions and /s/ in initial and medial positions. His Articulation Score would be computed in the following manner:

Sound	Initial	Medial	Final	Total
/r/	1.55	1.55	1.55	4.65
/l/	1.24	1.24	1.24	3.72
/s/	1.11	1.11		2.22
Total				10.59

His Articulation Score (AS) would be 10.59. It could be said that approximately 10½ per cent of his articulation was defective.

The Bell Laboratories sound frequency analysis considered the consonants as initial or final in a syllable. No medial sounds were recognized. The possibility that an Articulation Score based on the original figures rather than the prorated figures would be more accurate was considered. Therefore, a tabulation was made of each subject's articulation based on initial and final consonant sounds. The values for the vowels did not change. This new score was called AS' to differentiate it from the original AS (see Table VII).

In the previous example the subject misarticulated /r/, /l/, and initial and medial /s/. The test words for /r/ were ruler, carrot, and car. A consideration by syllables showed

TABLE VII

SOUND VALUES FOR THE ARTICULATION SCORE

Consonant Values Determined by Syllable Position (AS')

Sound		Values		
		Initial	Final	Total
/t /		2.52	6.49	9.01
/n /		1.44	3.69	5.13
/r /		0.80	3.85	4.65
/d /		1.81	2.40	4.21
/l /		1.24	2.48	3.72
/s /		1.93	1.39	3.32
/m /		1.70	1.62	3.32
/k /		1.76	1.27	3.03
/w /		2.70	----	2.70
/ð/	THen	1.94	0.37	2.31
/z /		0.10	1.94	2.04
/f /		1.32	0.79	2.11
/j /	Y	1.87	----	1.87
/h /		1.66	----	1.66
/v /		0.36	1.25	1.61
/b /		1.40	0.12	1.52
/p /		1.14	0.37	1.51
/g /	Gun	1.36	0.11	1.47
/ŋ /	NG	----	1.05	1.05
/ʃ/	SH	0.50	0.09	0.59
/θ /	THin	0.58	0.01	0.59
/tʃ/	CH	0.16	0.16	0.32
/hw/	WHen	0.26	----	0.26
/dʒ/	JudGe	0.24	0.04	0.28
/ʒ /	ZH	0.01	0.01	0.02
		28.80	29.50	58.30

TABLE VII (continued)

Vowels					
Key Word	Sound	Value	Key Word	Sound	Value
about	/ə/	6.80	pole	/oU//o/	1.98
pin	/I /	5.86	differ	/ɚ/	1.90
pen	/ɛ/	3.21	pawn	/ɔ/	1.73
pine	/aI/	3.16	pun	/ʌ/	1.73
pan	/ae/	2.87	pull	/U /	1.23
pot	/ɑ/	2.72	pout	/aU/	0.70
peel	/i /	2.69	purr	/ɝ/	0.33
pool	/u /	2.61	pew	/ju/	0.11
pane	/eI/ /e/	1.99	poise	/ɔI/	0.08
			Total		41.70

Initial Consonants	28.80
Final Consonants	29.50
Vowels	41.70
Total	100.00

initial /r/ in ruler and carrot and final /r/ in car. Syllable consideration of the test words for /l/ showed initial /l/ in lamb and umbrella and final /l/ in doll. The omitted /s/ in Santa Claus and pencils is initial in the syllable. The AS' would be computed in the following manner:

Sound	Initial	Final	Total
/r/	0.80	3.85	4.65
/l/	1.24	2.48	3.72
/s/	1.93		1.93
Total			10.30

The AS of this subject was 10.59; AS' for the same subject was 10.30. Further examples of the computation of AS and AS' will be found in the Appendix.

SUBJECTS

Forty-five children, all attending school in the Tucson area, between the ages of six and twelve were used for this investigation. The nine subjects attending the Tucson Public Schools were receiving therapy; the thirty-six attending Amphitheater Public Schools had received none. Two of the children were referred by teachers; all others were previously selected by speech therapists on the basis of the presence of articulatory defects. The articulation of all children ranged from normal to severely defective. None of their speech was unintelligible.

TEST PROCEDURE

Articulation Test. A picture articulation test constructed by the writer was used for testing articulation. All pictures were in color, and mounted on $5\frac{1}{2}$" by 7" pieces of poster board. Every sound in American speech was represented. The test was given to each child, and all sound errors were noted on the test forms. Copies of the articulation test forms will be found in the Appendix. Each child was given an AS and AS' by adding together the values of the sound errors.

Recorded Samples. Following the taking of the articulation test, each subject spoke until one minute of continuous speech had been tape recorded. The children talked about their Christmas presents or their pets. A few talked about brothers and sisters. All of the children were tested and recorded in school rooms.

RATINGS BY JUDGES

Morrison and others at Ohio State University found that trained judges could reliably place tape recorded samples of speech along a severity continuum.[30] In order to obtain a

---

[30] Morrison, Op. Cit.

Sherman and Morrison, Op. Cit.

March, et al., Op. Cit.

reliability _r_ of .90, ten judges and two speech samples for each subject or six judges and four samples were necessary.[31] All definitions of defective articulation relate the speech to the listener. It is therefore assumed that when a group of listeners can reliably evaluate the defective articulation, such evaluation is a valid measure of articulation defectiveness.

Training the Judges. Ten observers who were unfamiliar with the subjects participated in this study. All judges were upper division students in speech therapy at the University of Arizona and had clinical experience with articulatory defects.

Ten-second samples of the speech of forty-five children were compiled from the one-minute recorded segments of speech, which were used throughout the training and judging sessions. The same room and the same tape recorder were used for all sessions.

Half of the speech samples were presented at the first training session so that the observers could determine in their minds the range of defective articulation present on the tape. A tape was then played which presented ten-second samples of the speech of all subjects in random order. Judges were instructed to listen carefully to each sample of speech and to rate the articulation on a _one_ to _nine_ severity scale. _One_ was considered

---

[31] March, et al., _Op. Cit._

least severely and <u>nine</u> was considered most severely defective articulation. They were asked to consider only articulation, ignoring all influence of age or of general effectiveness. Each judge gave a rating to each of the forty-five speech samples on forms provided by the investigator (see Appendix).

All scores for the same child were averaged so that each child received a single score. The tape was played again, the composite ratings being announced before each selection was heard. If any observer felt the rating was not correct, that rating was discussed. For example, one boy received a better rating than some of the judges felt he should have. They thought that his apparent age, indicated by his voice and his subject matter (his teddy bear),, had influenced the ratings. The selections which were questioned were judged again, and new composite ratings were determined. This first training session lasted four hours.

The second training session, one hour in length, was devoted to listening to a tape of thirteen selections representing the various levels of articulation. The samples for this training tape were chosen in the following manner. The selections were first chosen upon which there had been close agreement by the judges during the first training session. These ratings were then compared with the AS and AS4, which had been divided into nine equal levels (see page 54). Those which appeared to be comparable were selected for the short training tape. The samples on the short training tape were arranged in descending order, one first and nine last. The judges were instructed to listen to this tape and to determine in their own minds the intervals which were shown.

Final Judging. The final judging session required two hours. The short training tape was presented twice at the beginning of the session so that the judges could determine the interval for each rating in his mind. A new tape, containing ten-second samples of the speech of all forty-five subjects in random order, was then played. After these samples were rated on the forms provided, the short training tape was again played. Different samples of the speech of the same children, arranged in random order, were then heard and rated by the judges.

It should be pointed out that there were originally sixty-five subjects in the study. In some cases the observers felt the sample was inadequate because of obvious mental retardation, poor voice quality, or poor recording conditions. All of those selections about which the majority of the judges indicated inadequate observations were eliminated.

CHAPTER IV

PRESENTATION OF DATA

Table VIII shows the Articulation Scores (AS) and (AS') and final judges' ratings for all subjects. The scores for AS ranged from 0 (normal) to 30.85. The range for AS' was 0 to 34.02. The ratings are on a <u>one</u> to <u>nine</u> scale, while the Articulation Scores ranged from 0 to 30+.

Table VIII shows that the scores for each child were similar whether the three consonant positions were considered (AS) or whether only the initial and final syllable positions were used (AS'). When a child made a sound error in one position, he usually made the same error in all positions. For this reason, the Articulation Scores were expected to be similar for each child.

The final rating each subject received was a compilation of the observations of all ten judges for both samples of speech. These ratings ranged from 1 to 8.4. The judges did not give fractional ratings; the fractions resulted when all ten ratings for both samples of speech were averaged.

Table IX shows the ratings and Articulation Scores for each subject in descending order. In order to compare ratings and scores for each subject, the ratings and scores were divided

TABLE VIII

JUDGES' RATINGS AND ARTICULATION SCORES

Subject	Rating	AS	AS'
1	6.25	20.51	20.83
2	1.70	6.56	6.57
3	2.00	6.56	6.57
4	1.60	3.93	3.91
5	1.25	3.43	3.48
6	8.45	30.85	34.02
7	7.95	20.31	21.40
8	6.75	23.15	22.69
9	6.40	19.14	21.06
10	4.30	20.57	19.32
11	7.40	29.68	33.25
12	1.70	5.56	5.56
13	2.35	6.56	6.57
14	2.55	6.56	6.57
15	1.80	6.17	6.53
16	4.95	16.57	16.55
17	3.05	10.41	10.78
18	3.45	13.22	13.18
19	3.55	17.60	19.82
20	6.25	17.25	17.99
21	1.90	5.37	5.36
22	1.20	3.33	3.32
23	1.65	8.18	7.97
24	2.20	6.16	5.95
25	1.30	4.40	3.74
26	1.20	0.00	0.00
27	2.30	5.57	5.37
28	3.35	16.19	17.11
29	3.00	8.64	8.65
30	3.45	7.80	9.05

TABLE VIII (continued)

Subject	Rating	AS	AS'
31	2.20	5.97	5.95
32	2.10	5.57	5.94
33	2.15	3.33	3.32
34	2.85	9.49	9.50
35	3.75	14.03	14.84
36	2.20	5.97	5.95
37	1.00	0.00	0.00
38	1.40	2.79	3.05
39	1.45	3.33	3.32
40	2.25	6.91	8.16
41	3.05	11.12	10.92
42	1.45	4.45	3.62
43	1.40	5.81	3.72
44	2.45	5.96	6.11
45	2.00	6.56	6.57

TABLE IX

JUDGES' RATINGS AND ARTICULATION SCORES

Rank Order

	Observations		AS		AS'	
	Subject	Rating	Subject	Score	Subject	Score
1	37	1.00	26	0.00	26	0.00
	22	1.20	37	0.00	37	0.00
	26	1.20	38	2.79	38	3.05
	5	1.25	22	3.33	22	3.32
	25	1.30	33	3.33	33	3.32
	38	1.40	39	3.33	39	3.32
	43	1.40	5	3.43	5	3.48
	39	1.45			42	3.62
	42	1.45			43	3.72
					25	3.74
2	4	1.60	4	3.93	4	3.91
	23	1.65	25	4.40	21	5.36
	2	1.70	42	4.45	27	5.37
	12	1.70	21	5.37	12	5.56
	15	1.80	12	5.56	32	5.94
	21	1.90	27	5.57	24	5.95
	3	2.00	32	5.57	31	5.95
	45	2.00	43	5.81	36	5.95
	32	2.10	44	5.96	44	6.11
	33	2.15	31	5.97	15	6.53
	24	2.20	36	5.97	2	6.57
	31	2.20	24	6.16	3	6.57
	36	2.20	15	6.17	13	6.57
	40	2.25	2	6.56	14	6.57
	27	2.30	3	6.56	45	6.57
	13	2.35	13	6.56		
	44	2.45	14	6.56		
			45	6.56		

TABLE IX (continued)

Observations			AS		AS'	
	Subject	Rating	Subject	Score	Subject	Score
3	14	2.55	40	6.91	23	7.97
	34	2.85	30	7.80	40	8.16
	29	3.00	23	8.18	29	8.65
	17	3.05	29	8.64	30	9.05
	41	3.05	34	9.49	34	9.50
	28	3.35	17	10.41	17	10.78
	18	3.45			41	10.92
	30	3.45				
4	19	3.55	41	11.12	18	13.18
	35	3.75	18	13.22	35	14.84
	10	4.30	35	14.03		
5	16	4.95	28	16.19	16	16.55
			16	16.57	28	17.11
			20	17.25	20	17.99
			19	17.60	10	19.32
					19	19.82
6	1	6.25	9	19.14	1	20.83
	20	6.25	7	20.31	9	21.06
	9	6.40	1	20.51	7	21.40
			10	20.57	8	22.69
7	8	6.75	8	23.15		
	11	7.40				
8	7	7.95	11	29.68		
	6	8.45				
9			6	30.85	11	33.25
					6	34.02

into nine groups. The judges rated the samples of speech according to the method of equal-appearing intervals; therefore, the Articulation Scores were divided into nine equal groups. The lowest AS was 30.85. Dividing 30.85 by 9 gave a group division point of approximately 3.43. Group <u>one</u> of AS consisted of those subjects whose scores fell between 0 and 3.43; group <u>two</u> consisted of those subjects whose scores were between 3.43 and 6.86 (3.43 + 3.43 = 6.86); group <u>three</u> consisted of scores between 6.86 and 10.29 (6.86 + 3.43 = 10.29).

The groups for AS' were not exactly the same as those for AS because the scores were not the same. Since 34.02, the lowest AS', divided by 9 is 3.78, group <u>one</u> fell between 0 and 3.78; group <u>two</u> fell between 3.78 and 7.56 (3.78 + 3.78 = 7.56); group <u>three</u> fell between 7.56 and 11.34 (7.56 + 3.78 = 11.34).

Judges' ratings between 1.00 and 1.45 were considered group <u>one</u>; ratings between 1.50 and 2.45 were group <u>two</u>; ratings between 2.50 and 3.45 made up group <u>three</u>.

The rank order of the scores and ratings for each subject is similar, and many of the same subjects can be found in any comparable group. The grouping of the subjects does not represent a normal distribution curve. All subjects in this study were considered to have defective articulation by teachers or speech therapists. In order to have a normal distribution curve, a random sampling of the entire population would have to taken. In this case, 5.0 would represent average, while ratings towards 1 would be better than average, and ratings towards 9 would be poorer than average. Judges' ratings in this study were relative to the samples presented; a rating of 1 represented least severely defective articulation.

The coefficient of correlation as computed from the raw scores produced an *r* of .9400 between the judges' ratings and the AS. The *r* for ratings of judges and AS' was .9413. These figures indicate that there are negligible differences between Articulation Scores when sounds are prorated equally in all positions in words or when the consonant sounds are given values determined by their position in the syllable.

CHAPTER V

SUMMARY AND CONCLUSIONS

Since this investigation was undertaken to construct an Articulation Score which would represent numerically a person's articulatory proficiency, it was necessary that the score be meaningful, accurate, and simple to use. The specific questions to be answered by the study were:

1. Can a scale be devised which will yield a single score which is numerically accurate, statistically manipulative, easily understood, and convenient and simple to use?
2. Is such a score valid as a representation of articulatory proficiency?

Review of the literature of the field of speech therapy revealed that the presently available methods of measuring articulation were inadequate. Most of the methods involved subjective evaluations of defectiveness by trained observers. The only numerical scale was based on the frequency of occurrence of consonant sounds in American speech.

The best way to construct a numerical scale of articulation seemed to be to assign values to each sound (consonants, vowels, and diphthongs) according to the number of times each sound would probably occur in a segment of American speech one hundred sounds in length. The Bell Laboratories sound frequency analysis was selected as the basis for the Articulation Score because it appeared to be the most carefully done of all the studies and because it was complete.[32]

Two Articulation Scores were constructed. Both scores were based on total values of the consonant sounds as determined at Bell Laboratories. The AS was constructed according to Wood's method--equally prorating the consonant sounds into all positions in which they could be found in American speech.[33] The vowel and diphthong values from the Bell Laboratories study were added to the total consonant values to equal 100 per cent. The AS' used the original Bell Laboratories distribution of consonant sounds according to their positions in syllables. The values for the vowels and diphthongs were the same in both scores.

Forty-five children between the ages of six and twelve served as subjects for the study. Their articulation ranged from normal to severely defective. Each subject was given a picture articulation test constructed by the writer. This test included all sounds of American speech.

---

[32] French, Carter, and Koenig, Op. Cit.

[33] Wood, "Parental Maladjustment and Functional Articulation Defects in Children," Op. Cit.

A one-minute sample of the speech of each child was tape recorded. Three ten-second samples were taken from the one-minute segments and arranged in random order. One sample of the speech of each subject was used to train the judges; the other two were used in the final judging.

Ten trained judges rated the samples on a one to nine continuum of articulation defectiveness, one meaning least severe and nine meaning most severe articulation defectiveness. A coefficient of correlation between each of the two Articulation Scores and the ratings of the judges produced an *r* of .9400 for AS and .9413 for AS'.

CONCLUSIONS

The results of this investigation allow several conclusions to be made:

1. A score can be devised which will numerically represent a person's articulatory proficiency.
2. The Articulation Score is a valid and reliable method of determining proficiency of articulation. The Bell Laboratories figures used as the basis for the score have been shown to be reliable[34] (cf. pages 36-38). The method of using trained judges to ascertain the severity of defectiveness of articu-

---

[34] French, Carter, and Koenig, *Op. Cit.*

articulation of given speech samples has also been shown to be reliable.[35] Since all definitions of defective articulation relate the defective speech to the judgment of a listener, the close correlation obtained between judges' ratings and the Articulation Score provides an indication that the AS and AS' are valid measures of articulation defectiveness.

3. Medial sounds may not be necessary in articulation tests and therapy. According to the results of this study, considering the syllable positions of consonants produces the same results as considering the position of the consonant sound in a word. It may be that a medial sound which is misarticulated or omitted is related to the initial or final position in the syllable rather than to its medial position in the word. This particular situation will have to be investigated further before a definite conclusion can be drawn.

DISCUSSION

The Articulation Score was constructed to fill a need in speech therapy. Although it was designed primarily for research

---

[35] March, et al., Op. Cit.

purposes, it can be used in clinical sutiations. Intelligence, attitudes, personality, aptitudes, adjustment can all be measured. All of these are at least as intangible as articulatory proficiency, and yet an assessment can be made.

Only children were used in the studies at Ohio State University.[36] Since their methods were used to validate the Articulation Score, it was necessary to use children in the present study also. However, even though this experiment was performed on children, the Articulation Score can be used for adults with equally satisfactory results. If both children and adults had been used, the possibility might have arisen of assessing the severity of defectiveness according to the "normal." A child of six may not be able to say /s/ and still be considered normal, while an adult should not misarticulate any sounds. The Articulation Score was not designed to tell how "normal" articulation is; its purpose is to designate the amount of speech which can probably be understood in ordinary conversation. For this reason, the age of the subject should not affect the validity of the score.

Since Morrison's rating scale used a descending order of severity, that is, one denoting least severely defective articulation and nine denoting most severely defective, the Articulation Score was also computed in the negative form. A score of 0 meant

---

[36] Morrison, Op. Cit.

that no articulatory errors occurred. It is suggested, however, that in actual use the total values of the misarticulated sounds be subtracted from 100 per cent, making the Articulation Score positive. A score of 100 would then indicate that all sounds were articulated correctly. The Articulation Score would then designate articulatory proficiency.

The comparison of the two methods of determining an Articulation Score--by prorating consonants equally in all word positions or by considering all consonants relative to their positions in syllables--was not an original purpose of this investigation. Wood's method of equally prorating the sounds[37] was the method selected for use. The results of this study indicate that either method results in similar scores. A definite conclusion concerning the two methods cannot be made at this time; many of the sounds misarticulated were consistent throughout all positions. A larger sample should be used so that the differences will be more apparent. The Articulation Score has been designed so that either table of values can be used. It cannot be stated which of the tables is more accurate until more research is done, determing the importance of medial sounds in articulation therapy and research.

---

37 Wood, "Parental Maladjustment and Functional Articulation Defects in Children," Op. Cit.

The Articulation Score has many advantages over any other method of articulatory measurement which has been developed. It is related to social adequacy of speech and is much easier to use than a subjective method of evaluation. It is easy to interpret and can be computed in a few minutes. The Articulation Score is a numerically accurate method of measurement which can be statistically manipulated. This means that measures of central tendency and tests for significant differences can be made from the Articulation Score. If a person can give an articulation test, he can use the Articulation Score. The mathematical operations are very simple. Reports to physicians, teachers, and other therapists (psychotherapists, physical therapists) can be made in a short period of time. As long as they know that the score represents articulatory proficiency, the score should be meaningful.

The Articulation Score is especially valuable for research purposes. It is discouraging to attempt articulatory research when a time-consuming method of determining articulation defectiveness, such as the ratings of judges, must be used. So much time must be spent getting defectiveness ratings that the original problem may be slighted. Articulatory defects are common, but research in this area is relatively rare. Only the causes can be studied, not the results of differing methods of therapy, unless progress can be assessed.

SUGGESTIONS FOR FURTHER STUDY

The Articulation Score opens the way for many research possibilities in the area of articulatory defects. The factor of maladjustment can be studied from a therapy point of view. For example, is there a difference in the results of therapy when the child also receives counseling as well as speech therapy, or when the parents receive psychological counseling while the child receives speech therapy? Whether psychological factors cause speech problems or whether speech problems cause psychological disturbances has not yet been determined. The problem can now be approached from a different point of view. Rather than spend time trying to decide the role of maladjustment in speech problems, study can be undertaken concerning the results of therapy when the psychological problems are also treated. In such cases it is not necessary to decide whether the maladjustment caused the speech problem or whether the speech problem caused the maladjustment. It may be that a circular problem has been set up--emotional tension causing functional articulatory disorder, in turn causing more emotional tension. Investigations from the point of view of therapy rather than etiology may help many people to get better and more lasting results from speech therapy.

In addition, other possibilities for study were brought to mind during the completion of this investigation. The judges who participated in the study expressed an interest in finding out whether their ratings would have been any different if they could have heard the children talk in person. They were especially concerned about the high frequency sounds, such as /s/, /f/, and /θ/. A comparison of these two methods of judgment ratings would be interesting.

Research concerning the importance of the medial sound in articulatory defects has been discussed. Although a numerical method of measuring articulatory profiency has been constructed, each individual who uses the Articulation Score must decide whether he prefers to use the original Bell Laboratories frequency of occurrence figures or whether he wishes to prorate consonant sounds equally in all positions, using the Bell Laboratories total values for each sound. The articulation test in a study of medial sounds should include medial sounds which are initial and final in the syllable, such as toaster, tomato, meatball, cat.

A suggestion was made that the Articulation Score be standardized so that a certain range of scores would mean mildly defective or severely defective articulation. No attempt at standardization was made in this study. As long as the score is not standardized, it can be used indiscriminately with any group of people. However, there are uses for such a standardization of the Articulation Score. Such a procedure would require a large, carefully selected population sample. Different age levels sould probably have to be studied individually. Judges would have to be trained to provide the comparison of the scores with a subjective evaluation of the amount of defectiveness present.

The writer hopes that the Articulation Score will open the way for thorough research in the area of articulatory defects.

BIBLIOGRAPHY

Anderson, Virgil A., Improving the Child's Speech. New York: Oxford University Press, 1953.

ASHA Committee on the Midcentury White House Conference, "Speech Disorders and Speech Correction." Journal of Speech and Hearing Disorders, XVII (June, 1952), 129-137.

Curry, Robert, et al., "A Phonographic Scale for Measurement of Defective Articulation." Journal of Speech Disorders, VIII (March, 1943), 123-126.

Dewey, Godfrey, Relativ Frequency of English Speech Sounds. London: Humphrey Milford, Oxford University Press. Cambridge: Harvard University Press, 1931.

French, Norman R., Charles W. Carter, Jr., and Walter Koenig, Jr., "The Words and Sounds of Telephone Conversations." Bell System Technical Journal, IX (April, 1930), 290-324.

Garrett, Henry E., Statistics in Psychology and Education, fourth edition. New York, Toronto, London: Longmans Green and Co, © 1953.

Henrickson, Ernest H., "An Analysis of Wood's Articulation Index." Journal of Speech and Hearing Disorders, XIII (September, 1948), 233-235.

Johnson, Wendell, et al., Speech Handicapped School Children. New York: Harper & Brothers Publishers, © 1948.

March, Nancy, et al., "Observed and Predicted Estimates of Reliability of a Speech Articulation Rating Scale." Speech Monographs, XXV (November, 1958), 296-304.

Mecham, Merlin J., A Scale for Measuring Level of Verbal Communication Behavior in Children. (Unpublished Work, Brigham Young University, Utah, 1958).

Morrison, Sheila, "Measuring the Severity of Articulation Defectiveness." Journal of Speech and Hearing Disorders, XX (December, 1955), 347-351.

Schneiderman, Norma, "A Study of the Relationship between Articulation Ability and Language Ability." Journal of Speech and Hearing Disorders, XX (December, 1955), 359-363.

Shames, George, "An Investigation of Prognosis and Evaluation in Speech Therapy." *Journal of Speech and Hearing Disorders*, XVII (December, 1952), 386-392.

Sherman, Dorothy, and Sheila Morrison, "Reliability of Individual Ratings of Severity of Defective Articulation." *Journal of Speech and Hearing Disorders*, XX (December, 1955), 352-358.

Travis, Lee Edward, *Speech Pathology*. New York, London: D. Appleton and Company, 1931.

Van Riper, C., *Speech Correction: Principles and Methods*, second edition. New York: Prentice-Hall, Inc., 1947.

Voelker, Charles H., "Technique for a Phonetic Frequency Count in Formal American Speech." *Extrait des Archives Neerlandaises de Phonetique Experimentale*, XI (1935), 69-72.

Wood, Kenneth Scott, "Measurement of Progress in the Correction of Articulatory Speech Defects." *Journal of Speech and Hearing Disorders*, XIV (June, 1949), 171-174.

________, "Parental Maladjustment and Functional Articulation Defects in Children." *Journal of Speech Disorders*, XI (December, 1946), 149-153.

APPENDIX

NAME ______________________ ARTICULATION SCORE 6.91 (8.16)

SCHOOL ______________ AGE 12 JUDGES' RATING 2.25

Sound	Check Words	I	M	F		AS	AS'
/t/	toaster, tomatoes, cat (kitty)(mew)	-			/l/	1.24	2.48
/n/	nest, Indian, lion				/s/	3.33	3.32
/r/	ruler, carrot, car				/z/	2.04	2.04
/d/	dog, candy, bed				/tʃ/	0.30	0.32
/l/	lamb, umbrella, doll			w		6.91	8.16
/m/	monkey, hammer, drum						
/s/	Santa Claus, pencils, house	t	t	t			
/k/	cake, bacon, clock						
/w/	watch, sidewalk						
/ð/	feather						
/z/	zebra, scissors, nose	d	d	d			
/j/	yellow, onion						
/f/	fork, elephant, knife						
/h/	horses						
/v/	vacuum cleaner, television, stove						
/b/	baby, rabbit (bunny), bathtub						
/p/	pipe, airplane, soup						
/g/	guns, wagon, pig						
/ŋ/	swinging						
/ʃ/	shoe, dishes, fish						
/θ/	thumb, toothbrush, teeth (mouth)						
/tʃ/	chicken, pitcher, match	t	?	?			
/dʒ/	jump, refrigerator, orange						
/hw/	whistle						
/ʒ/	television						

NAME ______________________ ARTICULATION SCORE 17.25 (17.99)

SCHOOL ______________ AGE 6 JUDGES' RATING 6.25

Sound	Test Words	I	M	F		AS	AS'
/t/	toaster, tomatoes, cat (kitty)(mew)	-			/t/	3.00	2.52
/n/	nest, Indian, lion				/r/	4.65	4.65
/r/	ruler, carrot, car	w	w	w	/s/	3.33	3.32
/d/	dog, candy, bed				/ʃ/	0.77	1.94
/l/	lamb, umbrella, doll				/z/	2.04	2.04
/m/	monkey, hammer, drum				/θ/	0.60	0.59
/s/	Santa Claus, pencils, house	-	-	-	/tʃ/	0.10	0.16
/k/	cake, bacon, clock				/dʒ/	0.27	0.28
/w/	watch, sidewalk				/hw/	0.26	0.26
/ð/	feather		d		/ɚ/	1.90	1.90
/z/	zebra, scissors, nose	d	d	d	/ɝ/	0.33	0.33
/j/	yellow, onion						
						17.25	17.99
/f/	fork, elephant, knife						
/h/	horses						
/v/	vacuum cleaner, television, stove						
/b/	baby, rabbit (bunny), bathtub						
/p/	pipe, airplane, soup						
/g/	guns, wagon, pig						
/ŋ/	swinging						
/ʃ/	shoe, dishes, fish						
/θ/	thumb, toothbrush, teeth (mouth)	f	f	f			
/tʃ/	chicken, pitcher, match			?			
/dʒ/	jump, refrigerator, orange	d	d	d			
/hw/	whistle						
/ʒ/	television					ɚ, ɝ	

JUDGE ____________________

CHILD	RATING	CHILD	RATING	CHILD	RATING	CHILD	RATING	CHILD	RATING
1		27		53		79		105	
2		28		54		80		106	
3		29		55		81		107	
4		30		56		82		108	
5		31		57		83		109	
6		32		58		84		110	
7		33		59		85		111	
8		34		60		86		112	
9		35		61		87		113	
10		36		62		88		114	
11		37		63		89		115	
12		38		64		90		116	
13		39		65		91		117	
14		40		66		92		118	
15		41		67		93		119	
16		42		68		94		120	
17		43		69		95		121	
18		44		70		96		122	
19		45		71		97		123	
20		46		72		98		124	
21		47		73		99		125	
22		48		74		100		126	
23		49		75		101		127	
24		50		76		102		128	
25		51		77		103		129	
26		52		78		104		130	