THE JOURNAL OF SPEECH AND HEARING DISORDERS

Monograph Supplement 4

December 1954

PUBLISHED QUARTERLY BY THE AMERICAN SPEECH AND HEARING ASSOCIATION

Monograph Supplement 4, 1954

The Journal of Speech and Hearing Disorders

THE DISORDER OF ARTICULATION: A SYSTEMATIC CLINICAL AND EXPERIMENTAL APPROACH

Robert Milisen and Associates

Speech and Hearing Clinic
Indiana University

MANUSCRIPTS and related correspondence should be addressed to: Grant Fairbanks, Editor, *Journal of Speech and Hearing Disorders*, 321 Illini Hall, University of Illinois, Urbana, Illinois.

SUBSCRIPTIONS and orders for back numbers should be addressed to: George A. Kopp, Business Manager, *Journal of Speech and Hearing Disorders*, Speech Clinic, Wayne University, Detroit, Michigan.

NEWS and announcements should be addressed to: Charles R. Elliott, School of Speech, Northwestern University, Evanston, Illinois.

NOTICE TO AUTHORS

Before submitting manuscripts for publication authors should consult Information for Contributors to the Journal of Speech and Hearing Disorders, *JSHD*, 14, 1949, 93-94. Prospective authors are invited to write the Editor for copies of this Note, and of its supplement, Examples of Manuscript Form.

TABLE OF CONTENTS

FOREWORD

THE DISORDER of articulation is the subject matter of Monograph Supplement 4. The statement is straightforward. It begins with a theory, discusses its implications and then reports a group of experiments. There is much here to increase our understanding of the disorder and improve our techniques for its management. There is also nourishment for our hope that the most common of the speech disorders may not always be the most neglected.

The senior author, Professor Robert Milisen received his Ph.D. degree under the direction of Professor Lee Edward Travis in 1937 at the State University of Iowa, and has been Director of the Speech and Hearing Clinic at Indiana University for many years. All of his associates completed their M.A. degrees under Professor Milisen's guidance, all at Indiana and all in 1953 and 1954. Four of them are actively engaged in speech therapy, William R. Humphrey with the Bay State Society for the Crippled and Handicapped in Worcester, Massachusetts, Donald B. Rice, Edward F. Romans and Davis A. Scott in the public schools of Bloomington, Indiana, Chicago, Illinois and Sioux City, Iowa, respectively. Katherine Snow has remained at Indiana University as Instructor in Speech and Hearing Therapy. Herbert N. Wright is pursuing advanced graduate study at Northwestern University, where is a graduate assistant.

The Editor wishes to acknowledge the collaboration of Associate Editor Claude E. Kantner, and to thank Assistant to the Editor Elaine Pagel Paden, Marchetta Bowers and Ramona J. Smith for their labors in steering the manuscript into print.

Grant Fairbanks
1 December 1954

1

A RATIONALE FOR ARTICULATION DISORDERS

Robert Milisen

IN APPLIED sciences, practical procedures should be tested experimentally whenever possible. In an applied field of a clinical nature, unfortunately, testable basic concepts may not have been structured and symptoms may be studied and treated without regard for their relation to the whole problem. This is likely to be the case in much articulation therapy.

The purpose of this rationale is to demonstrate briefly a *Basic Concept* of articulation relating the conditions precipitating and maintaining the disorder to examination, and diagnostic and therapeutic methods of a practical nature which can be measured experimentally. Older sources were reorganized along with new material to show the need for and the means of diagnosing the child's articulatory *skills*, as well as his liabilities, to enable therapy to cope best with each child's problems.

The basic concept hypothesized is that defective articulation, a substitute response for normal articulation, results from the disruption of the normal learning process. Also, this defective response need not be maintained in any child provided he receives appropriate training which builds on his skills, begins early and continues long enough.[1] Furthermore, articulation therapy should be composed of principles found in the normal learning process, which should be directed positively toward production of the correct responses rather than negatively toward 'unlearning' of the defective responses. Lastly, therapy must begin at the level at which the normal process was disrupted in order to build on articulatory skills which the child already possesses.

This basic concept presupposes the existence of a means of describing and diagnosing the most advanced level of articulatory skill for each child as well as a therapy which can be readily adapted to each child.

Since it is vital for the diagnosis and therapy to be closely correlated with the normal learning process it becomes necessary to describe this process, as well as the ways in which it may be disrupted.

Normal Learning Process

The motor skills of normal articulation are phenomenally intricate and rapid. They grow out of the infant's simple movements and increase sufficiently in number, ease of production and complexity only if they are useful in communicating the infant's needs to the environment.

The skill of articulation has its roots in these movements and sounds which are made thousands of times during the first year. The movements and sounds beginning at birth are largely the mass response of an unhappy organism. From this mass are developed many separate movements and sounds as the infant grows older. The increase in frequency and the differentiation of movements and sounds are partly due to physical growth of the child but largely to his interactions with the environment which have reinforced his movements or sounds. Movements will not be repeated frequently and further differentiated if they are not reinforced. The reinforcement may come from the infant himself as he is becoming

[1]This is true except for those children who have severe flaccid paralysis of the speech mechanism or those who are from a neurological standpoint essentially decorticate organisms.

aware of his body movements and noises. The reinforcements may also come from the environment when it responds to these movements and noises which have become cues to the infant's needs. The more nearly the environmental behavior satisfies the infant's needs, the more likelihood that the movements and noises which elicited the response will be repeated. As new needs arise, new movements and sounds will differentiate from earlier ones and those will be repeated which are followed by an appropriate environmental response.

During earliest infancy the movements and sounds are more reflex in nature and are made in response to vegetative needs; nevertheless, they begin to create a pattern of communication. As the movements and noises become more 'voluntary,' such as in vocal play, the movements are practiced many times through the infant's self reinforcement and many more times through environmental reinforcement. This begins to develop a pattern of communication not based solely on vegetative needs of the organism but on his need to interact with the world around him. At first, attention to his movements and noises will suffice, but presently more specific responses are needed because the infant's needs have become more complex. The effort to obtain the desired response results in the emergence of new movements and noises as well as in a keener awareness of the environment. Hence, environmental noises which were not used in vocal play are imitated. This ability to imitate new sounds plus the need for a more specific medium of communication plus the willingness of the environment to provide training leads to the first word. The learning process which started with the interaction produced by the mass activity at birth continues until the child uses complex speech with adequate articulation.

The breakdown in this normal learning process results in the substitution of other kinds of learning, such as defective articulation, and may occur at any time and for a multitude of reasons. Any condition involving the environment (the teacher) and the infant (the learner) which interferes with the communication attitude will disrupt the normal learning process and produce substitute responses. When the learning process is disrupted, sounds and movements ordinarily expected are not repeated and strengthened, and the communication attitude leading toward adequate speech is neglected or destroyed.

A sound will not be repeated frequently unless it is reinforced. As a result, that sound or movement will be awkward if it is used later on in speech because motor skills for it were not developed during the pre-speech period.

Sounds and movements which are not reinforced will not create an attitude of communication and may create a stable withdrawal pattern. A communication attitude is necessary in stimulating the infant to differentiate new sounds from his old ones, and to imitate entirely new sounds and movements with which he can communicate his needs more specifically. The development of a withdrawal pattern limits the number of sounds and movements a child learns to make during his pre-speech period. The fewer sounds he learns to make the fewer words will he be able to articulate understandably when he begins to speak; hence, the less frequently will he be able to communicate his needs by speaking and therefore the less frequently will his speech be reinforced. The presence of a withdrawal pattern delays his effort to speak because he uses a different method in achieving adjustment to his environment. Attempts to force the 'delayed' child to speak often result in failure, since he has neither the motor skills needed for speech nor the desire to communicate through speech. This frequently leads to negativism because the environment expects the impossible of the child and punishes him for 'failing.'

Conditions which precipitate and maintain articulation defects after the child has begun to speak are only an extension of the conditions which limited the production and differentiation of sounds and which interferred with the development of a communication attitude before he began to speak. The first word spoken meaningfully is really a step toward the end of the learning process instead of the beginning.

To rephrase the basic concept as it relates to the precipitation and maintenance of articulation disorders: It makes no difference whether the infant's failure to develop the skills and attitudes necessary for speech with good articulation was due largely to his limitations or those of the environment, the difficulties could have been overcome and the child could have had adequate articulation if the environment had been trained to begin early in creating a desire as well as a medium of communication.

Many persons will react violently to this last statement by citing a variety of 'organic' cases which couldn't be 'cured.' The chief answer is that speech therapy usually begins long after most of the learning period of the infant's life has been wasted. His movements and sounds precipitated by vegetative needs were not built upon, hence a medium for and desire to communication were not created. Worst of all, the infant learned to accept a vegetative status, an attitude which is difficult and frequently impossible to break down. Further rationale for the statement that 'organic' cases do not disprove the hypothesis are found in the nature of the speech of people who have organic disabilities. For instance, children and adults with cleft palates or any other physical limitation do not have consistent types of misarticulation which can be diagnosed by studying the physical disability itself. Also children and adults with any and all types of physical disability have on occasion achieved adequate articulation, frequently without professional help.

This concept that misarticulation is a substitute response precipitated by a breakdown in the normal learning process should simplify articulation therapy because one does not need to create distinct types of therapy for all of the 'organic' and 'non-organic' groups. He will instead deal with each case on its own merits which will be determined primarily by behavior and speech performance, not by appearance and inheritance. At the same time the therapist will be able to adjust to the limitations of the environment and/or the child by placing proper stress upon various aspects of the diagnostic and therapeutic procedures.

In a clinical situation, therapy must be an outgrowth from the examination and diagnostic procedures since they should describe the nature of the disorder, and also demonstrate where in the learning process to begin the therapy as well as the kind of therapy to use.

In this article these conditions will be reversed in order to aid the reader. The therapy and its concepts will be developed first and then the methods of examination and diagnosis. The reader can then determine whether the examination and diagnostic methods will provide the information needed in setting up a therapy program.

Therapy Program

The program must begin at the most advanced level of articulation skill which the child can perform. New skills to be learned must build on an established one in such a way as to produce success quickly and easily.

The criteria which the therapy must fulfill are as follows:

1. It must cope with articulation deficiencies of all kinds, people of all ages, and of all environmental and physical limitations.

2. It must involve teaching methods, based on learning theory, which are easy to understand, easy to teach, able to provide the defective person with successful performance of those speech responses which are usually formed defectively, and to challenge him to greater achievements.

3. It must present the speech material to be elicited, practiced, and automatized in the form of a whole speech response (a sound or a more complex speech unit). It must be presented in such a manner and at such a level of articulation skill as to enable the child to produce the unit correctly (at least part of the time). For only a few sounds misarticulated by a few children, will it be impossible to begin with a whole speech unit. In these cases the dangerous fractionalizing methods must be used with care.

4. The scope of the therapy, its concepts and methods must not conflict with psychotherapy. (It is true that in addition to his articulation problems, the child may have psychological problems, as do most people. Granted that these additional problems may need to be dealt with, especially in the older cases, but the therapist should not call this *articulation therapy*.)

The therapy method which fulfills the above criteria may be called *Integral Stimulation*. This requires *vividness in every aspect of stimulation* in order to make up for the child's deficient skills and therefore to make the 'sound movement' easier to imitate. The reason for changing to the term *Integral Stimulation* from the traditional 'Stimulation' or 'Stimulus Method' is to stress the fact that this is a much more vigorous and complete stimulation than the weak, incomplete, aimless kind given by so many clinicians. Integral Stimulation means whole stimulation and makes use of all the stimulation complex needed to elicit a correct response. It is the basic method used in eliciting practically all misarticulated sounds and therefore is used at the beginning of most therapy programs. 'Stimulation' in the hands of many clinicians does not elicit a correct response. These clinicians are then forced to revert to one of the fractionalizing methods which are usually not as adequate and are frequently dangerous because the residual movements not necessary to speech are reinforced unintentionally and therefore become confused with the sound to be learned. These residual movements are frequently retained.

Each Integral Stimulation includes the stimulation of the child with a whole speech configuration, and his imitation of it. It is divided into three parts: (1) the production of the sound by the clinician so the child hears and sees and perhaps feels it, (2) the child's response, which he and the clinician hear and see and feel, and (3) the evaluation of the response by both the child and the clinician.

The sound, when produced by the clinician, must emphasize for the child: (1) the way the sound sounds, (2) the way the clinician looks while producing the sound with special effort to make the focal articulation point visible, (3) the way the sound feels which facilitates for some children the production of some sounds, (4) the evaluation of the audible and visible characteristics of the child's imitated sounds by the clinician and also by the child, and (5) motivation to correct speech, a motivation which arises from the additional confidence created by immediate improvement from the Integral Stimulation.

The importance of vividness, forcefulness, and completeness in this stimulation cannot be over-stressed. The reproduction of the sound by the child is dependent largely upon the effectiveness of the stimulation. *It must be remembered that a speech movement skill cannot be learned until it is produced.*

Integral Stimulation therapy is based on three assumptions:

1. A child is most strongly motivated to correct his misarticulations if he can see and hear himself producing the whole sound correctly.

2. At the outset of therapy some misarticulated sounds in each child can be improved by this type of stimulation.

3. Common elements learned while correcting easily stimulated sounds are transferred to the sounds which were not easily stimulated in the beginning and will therefore cause them to become more stimulatable.

This enables the clinician to use the Integral Stimulation method with most of these 'difficult' sounds without reverting to procedures which fractionalize the speech process such as the placement method and oral calesthenics.

The Integral Stimulation portion of the therapy, in brief, involves (1) beginning with misarticulated sounds which respond well to stimulation. The child's skill in articulation of these sounds is facilitated by Integral Stimulation until he is able to produce them correctly and automatically. During this time, rapport is established on the basis of successful learning, and motivation increases as the child becomes acquainted with and confident in the stimulation method of therapy. (2) The clinician can then begin work with sounds which previously were not stimulatable, but which, because of transfer of training of common elements from the recently learned sounds, have become more stimulatable. (3) Any remaining sounds which have not improved are stimulated vigorously in an effort to produce any kind of a change which, through proper reinforcement, can be modified until it is developed into a correct imitation of the sound being stimulated. (4) If any sounds remain which still do not respond to stimulation, the clinician may have to regress to ear training and, if this does not bring success when followed up by stimulation, to one of the fractionalizing methods.

Although the eliciting of the correct sound through Integral Stimulation is the beginning of the articulation therapy, the learning process leading to good articulation involves the manipulation of many other variables.

The failure to organize therapy

around these variables, which are related to each child's skill in articulation, will at least slow down rehabilitation and frequently result in failure because the therapy was not adjusted to the proper level of difficulty. Hence, one child may be given a task which is so easy that it presents no challenge while another receives a task so difficult that he can not perform it.

Each of these variables can function independently to influence the level of difficulty of the speech performance. The wise therapist will keep two variables constant while manipulating the third. These variables are:

1. Amount of assistance needed by the child to produce correctly the speech configuration.[2]

2. The complexity of the speech configuration to be produced.

3. The audience threat to good speech.

The first variable involves the amount of assistance needed by the child in producing correctly any speech configuration. It is influenced primarily by the amount of learning which has already taken place and varies from sound to sound. This variable may be subdivided into three parts:

a. Eliciting the sound by direct assistance to the child before he makes the sound. This can usually be accomplished by Integral Stimulation. This is a comparatively simple stimulus-response pattern. It does not require too much 'thinking' on the part of the child. It should never be threatening to the child if the proper sounds are stimulated. The proper sounds are those which the child can modify after receiving stimulation.

It must be remembered that any successful imitation of the therapist's stimulation indicates that the child has an adequate speech mechanism, enough alertness, and has learned enough auditory discrimination and coordination to make the sound, provided he has a crutch to lean on: Integral Stimulation. With additional therapy, the child's responses become strengthened, and he learns to make the sound without the 'crutch.'

b. Eliciting correct sound in any configuration by directing attention to the process, not to the content, of speech. This results in controlled speech. The variable at this level is such that the child can make the sound correctly and without stimulation but only when he 'thinks' of *how* he is speaking and not so much of *what* he is saying. This disrupts the faulty but habitual pattern and allows the child to substitute the new but less automatic speech pattern in its place. Much self evaluation must be used to help the child past this level, and he should work with only *one* sound at a time.

c. Eliciting automatic speech by directing attention to the content. Articulation skills newly learned in the controlled speech stage break down when the child speaks automatically. These errors can be corrected by eliciting automatic speech for a period of time and afterward directing attention to the misarticulations of the newly learned sounds. The additional attention directed to the process will enable the child to produce the sound correctly and strengthen his articulation skill, thereby increasing the likelihood that the sounds will be produced correctly even when spoken automatically. Presently, the sound will be used cor-

[2] A speech configuration is a unit of speech of any complexity which is at least as complete as a speech sound in isolation.

rectly and automatically in all situations.

The second variable involves the effect the complexity of the speech configuration will have on the accuracy of articulation. As a rule, a sound in isolation is easier to produce than one in a nonsense syllable, in a word, or in a sentence. Usually, a consonant adjacent to a vowel is easier to produce than one blended with another consonant. Memorized material is usually easier to produce than that which is read and much easier than that produced 'conversationally.' These generalizations are not always correct and only properly administered and evaluated tests will point out the deviations of each child from the general rule. Nevertheless, it is obvious that a child with severely limited articulation skills should be given speech configurations which are 'easy' enough for him to produce while the child in an advanced stage of learning should be stimulated with more complex configurations.

The third variable involves the effect the audience has on the adequacy of articulation. This effect is so powerful that it ranges from facilitation to complete disruption. A friendly audience usually facilitates a child's performance. An exception occurs if a friendly audience has previously accepted and reinforced faulty speech. A child in the presence of such an audience may regress to an earlier level of speech as a means of pleasing the audience and of obtaining attention. An unfriendly audience may present such a threat that the child regresses catastrophically to a faulty but more automatic speech pattern.

Each of these variables can be further subdivided and described. But, in this paper, suffice it to say that articulation therapy to be successful must be dynamic enough to change constantly so as to keep the interrelationships between these variables such that the level of difficulty for the child will always challenge him to articulate better but never leave him feeling a complete failure. The ultimate goal is to enable the child to articulate correctly under the most advanced condition of all three variables.

Examination and Diagnosis

In planning therapy for a child with defective articulation, it is important to have information about precipitating conditions, physical and mental capacities, emotional problems and the child's home environment, in addition to detailed knowledge of his articulation status and his stimulability. Sources of this information are the case history, mental and physical tests and special disability tests which need not be described here because they have been described elsewhere. The articulation and stimulation tests will be described because some variation is presented from the traditional testing techniques.

First, it is desirable to attempt to assess the general nature of the misarticulations without trying to record them. *A Conversation-Articulation Test* involving pictorial material, toys, pets, etc. may be used. Conversation directed toward the parents is sometimes helpful in precipitating speech from the child. This makes it possible to determine whether the child comprehends speech, whether he can speak understandably, and what his attitude and that of his parents is toward his speech.

Easy and interesting material for an *Oral Reading-Articulation Test* will allow comparisons between articulation while conversing and reading and may allow observation of sounds not heard during the conversation test.

Material which the child has memorized, such as poems and songs, can sometimes be used as a *Memory Material-Articulation Test,* particularly when parents are present and can help the child to start. Songs may be useful as a testing device with the very timid child if the parents will join in.

The chief examination, the *Isolated Word-Articulation Test,* permits the examiner to record the nature of the speech responses for each sound in each position, under constant conditions. The sound in a spoken word should be elicited with pictures for young children and word lists for older children and adults. It is important that the sounds be testable in each position in which they occur, that the errors be recorded in sufficient detail to be useful to the therapist, that the evaluations of a given defective sound be consistent among examiners, that the words used be in the vocabulary of the child and that consonant sounds tested be adjacent to vowels except when blends are intentionally used for testing purposes.

The method of recording errors differs from the traditional pattern especially for indistinct sounds. The response to every sound tested is recorded as follows: omission–'om'; substitution–the phonetic symbol for the sound produced; a severely distorted sound which would attract the attention of most untrained observers and frequently make the sound difficult to recognize–'3'; a mildly distorted sound which would attract the attention of many untrained listeners but would not make it difficult to recognize–'2'; and a sound produced correctly–'1.'

The data sheet used in the *Isolated Word-Articulation Test* is much the same as the traditional ones with the following exceptions: (1) The sounds in the left column are represented by phonetic symbols which are arranged so the most commonly misarticulated ones are listed toward the top. (2) A number representing the frequency with which each sound occurs in the language is placed next to each phonetic symbol. (3) Three additional columns are added for the Isolated Word Test and seven additional columns are added for the Stimulation Test. A sample of the data sheet is given in Table 1.

TABLE 1. Sample of the data sheet used in the Isolated Word-Articulation Test.

Sound to be tested and its frequency	*Isolated Word-Articulation Test*			*Stimulability Test*						
				Sound in isolation	*Sound in nonsense syllable*			*Sound in word*		
	Initial	*Medial*	*Final*		*Initial*	*Medial*	*Final*	*Initial*	*Medial*	*Final*
r – (3)										
l – (5)										
[illegible] – (4)										
z – (9)										

The administration of the test requires the examiner to keep the child so interested that speech responses are made *spontaneously* and without embarrassment. For the most part the examiner should not talk about 'testing' the child. He should not administer the entire battery of sounds to all children since the great majority of cases will not misarticulate sounds past the first 15 of the data sheet. The informal tests mentioned earlier involving conversation, oral reading and memory should provide the cues as to the consistency and severity of the misarticulations.

Stimulability Test. After the articulation tests are completed, all misarticulated sounds are subjected to Integral Stimulation and at varying degrees of complexity of the speech configuration, that is, the sound in isolation, in nonsense syllables, and in words. Articulation responses following stimulation are recorded in the same manner as on the *Isolated Word-Articulation Test.* The examiner compares these results of the *Stimulability Test* with those obtained on the four articulation tests and thereby determines the effect of stimulation upon the misarticulated sounds. Children are able to improve as many as 85 per cent of their misarticulated sounds after receiving only a *few* integral stimulations. Any test which measures this ability would be invaluable to therapy.

The *Stimulability Test* was designed to provide an index of articulation skill by testing, evaluating and recording the quantitative and qualitative changes in articulation which follow a few Integral Stimulations. This index records the changes occurring in response to stimulation for each misarticulated sound in each position and on three levels of speech configuration—isolation, nonsense syllables, and words. This index provides concrete evidence as to which sounds respond immediately and well to Integral Stimulation, which ones are only modified, and which ones are not changed at all. Therefore, the results from this test will pinpoint the sounds on which the therapy should begin as well as the kind of therapy which will be most likely to precipitate improvement through use of Integral Stimulation.

The administration of the test involves directing the child's attention to the examiner's face by saying, 'I am going to make some noises. You *watch* me and *listen* and do what I do.' The examiner makes the sound, syllable, or word two or three times before signaling the child to imitate. This procedure is repeated twice and the best of the child's responses is recorded. The clinician must use Integral Stimulation by making the focal articulation point more visible, by giving clues as to the movement of breath stream, by urging the child to 'watch' or 'listen' carefully if the child's attention is straying and by a wise use of *praise.* The examiner should always begin by stimulating a sound which the child can reproduce correctly in order that praise can be given thereby reducing the threat of the testing situation. The examiner should always praise the response if the child has shown improvement. He may make qualified statements of praise such as, 'That is better,' if the imitation is better than the child's previous one. He should never say, 'That is wrong,' even though the imitation was completely incorrect. Instead, he should say nothing or casually say, 'Let's try it again.'

After the *Stimulability Test* is completed, the examiner should list each misarticulated sound in each of two columns. The sounds in the first column should be in such an order that those most distracting to the audience are toward the top. The same sounds would be listed in the second column with those most stimulatable toward the top. The obvious goal of therapy would be, as much as possible, to work first with those sounds which are toward the top in both columns because they would be the sounds which were most easily stimulated yet at the same time are the ones most distracting to the listener.

In the first column the sounds are listed toward the top which are most distracting to the listener. The degree of distraction is determined by combining quantitative and qualitative factors of each defective sound. That is, the frequency with which a misarticulated sound occurs (numerical rating is given in the first column of the *Isolated Word-Articulation Test* sheet) is combined with the severity of the misarticulation occurring each time the sound is produced. Thus, a sound which occurs frequently and is misarticulated badly each time it is produced would be placed toward the top of the column. A sound which does not occur frequently and is misarticulated mildly would be placed toward the bottom of the column.

The primary factors determining the order in the second column (Ease of Stimulation) are the responses on the *Stimulability Test* compared especially with the responses of the *Isolated Word-Articulation Test.* Sounds are considered more stimulatable if they are imitated better in words than in nonsense syllables, or in nonsense syllables than in isolation. Sounds are considered more stimulatable if the quality of the misarticulation is a mild distortion, rather than a severe distortion, or a severe distortion rather than a substitution, or a substitution rather than an omission. Other factors to be considered are the ease or struggle of the child during the attempted response and whether the stimulation produces some change even if not a correct one.

The secondary factors involved in estimating the probable effectiveness of Integral Stimulation Therapy are visibility of the focal articulation point, auditory acuity and discrimination, general speech environment, organic conditions and motivational factors. Since the child is more likely to respond to stimulation that he can see and hear, the therapist should use every device to make all sounds heard and all focal articulation points visible. The latter is especially important with those sounds which when produced in the ordinary way are only partially visible or are not visible at all.

Sounds made defectively by family, friends, or teachers may be more difficult to correct since the child is exposed to a lot of faulty stimulation. There may also be resentment by the adults who assume that the child is being taught *fancy* speech.

Sounds using parts of the speech mechanism which are defective, or poorly coordinated muscles may be more difficult to produce.

Motivation may be more important in correcting some sounds than others since the importance of a misarticulation to a child frequently determines how hard he will try. He may, for instance, be more willing to work on a misarticulated sound in his name than on any other misarticulated sound.

By proper integration of the primary factors with the secondary factors, the diagnostician can list the misarticulated sounds in order, with the most stimulatable sounds on top.

After the two lists are completed, the diagnostician is in a position to determine which sounds should be given therapy first and the methods to be used. The ultimate outcome of articulation therapy is often determined by the results which are obtained when treating the first sounds. The correction of a misarticulated sound not only improves the quality of the speech, but it also produces constructive changes in attitude in the child and his environment. This facilitates the therapist's efforts when dealing with the remaining sounds. Hence, one sees the vital importance of treating a sound that can be easily corrected, thereby establishing for the child and his environment a learning pattern as well as the motivation to follow it until speech is 'normal.'

The obligation of the diagnostician is to furnish methods of therapy which will produce normal speech in the shortest period of time without creating emotional problems in the child or in his parents. This is comparatively easy in most children because the misarticulations are neither caused by maladjustment nor have they precipitated it to any marked degree. By beginning therapy with the most distracting sound which is readily stimulated and building a stimulation therapy around the three previously mentioned variables, it should be possible to produce enough improvement in almost all cases to keep the child and the parents motivated until the speech is normal.

Many children having misarticulations also have maladjustments which may or may not arise from the defective speech. If the adjustment problem arose from the defective speech, it may implement or depress the child's efforts to achieve better speech. If it implements therapy, the child will, although bothered by his disorder, want to *try hard* to correct it.

The first sound to be corrected for such a child should be the one most distracting which is also stimulatable. The continuance of the high motivation of the child will depend on his awareness of speech improvement. The maladjustment will usually disappear as soon as the speech becomes adequate.

If the maladjustment results in a withdrawal from speech therapy, the diagnostician must choose the misarticulated sound most easily corrected regardless of how much or how little it distracts the listener. The goal at the outset is to build confidence for the child in his ability to improve his articulation and in the method of therapy. Success on the first sound leads to a willingness to 'try' the others.

If the maladjustment has not arisen from the defective speech, the diagnostician should provide for psychotherapy as well as articulation therapy. If psychotherapy is not available, articulation therapy can be started providing Integral Stimulation produces improvement and provided the child does not use his defective speech as an attention seeking device to compensate for his other maladjustments. Speech improvement in a maladjusted child, like improvement in any other skill, will produce some self confidence which may help to compensate for the non-speech maladjustments. Therapy which in a natural manner produces improvements will not create new maladjustments nor cause old ones to get worse.

Experimental Evaluation

Many assumptions made in this Rationale have been verified clinically, but few have been evaluated experimentally. An effort, however, has been made to subject to experimentation the following assumptions, as reported in the succeeding articles:

1. The articulation responses to the Isolated Word-Articulation Test and the Stimulatability Test, although divided into seven classifications, can be evaluated reliably by different examiners. (Herbert Wright)

2. Words in the Isolated Word-Articulation Test should be elicited by pictorial rather than oral methods. (Katherine Snow)

3. The effectiveness of various aspects of integral stimulation differs from one sound to another and from one child to another. Partial stimulation is seldom as effective a means of eliciting a sound as is the total or integral stimulation. (Davis Scott)

4. Prognosis for a misarticulated sound can be obtained by contrasting the results obtained from pictorial and oral methods of eliciting those sounds. (Katherine Snow)

5. Integral Stimulation will implement the production of most misarticulated sounds. The amount of implementation is related to the nature of the stimulating conditions and the complexity of the speech pattern. (Davis Scott)

6. Articulation therapy being primarily a problem of teaching the production of a sound can be studied by measuring the behavior of normal speakers who are learning to produce unfamiliar sounds. The design used in this study should provide a vehicle for subjecting the articulation learning process to carefully controlled experimentation. (William Humphrey)

7. Many conditions involving integral stimulation will affect articulation therapy. One of those factors is the latency between stimulation and response and its effect upon the adequacy of the responses. A design similar to Humphrey's was used. (Edward Romans)

8. Another factor affecting integral stimulation therapy is the relation of the amount of stimulation to the degree of learning and retention. A design similar to Humphrey's was used. (Donald Rice)

Only the surface has been scratched in the study of defective articulation. Perhaps this Rationale and the experimental approach it provides will benefit the therapist as well as the experimenter.

II

RELIABILITY OF EVALUATIONS DURING BASIC ARTICULATION AND STIMULATION TESTING

Herbert N. Wright

IN APPROACHING the nature of defective articulation, much work has been done by correlating defective articulation with mental age, auditory memory span, general muscular coordination and other variables assumed to be related to articulatory problems. These studies are valuable, but there has been a paucity of research concerned with the actual measurement of the articulation of speech sounds, *per se;* and in earlier research studies adequate reliability of the examiners has been assumed. The studies by Henderson (*3*), Irwin and Chen (*4*), and Irwin and Curry (*5*) indicated that the articulation of a particular sound may be reliably measured as to the specific sound articulated or as to the correctness or incorrectness of the sound.

Curry, Kennedy, Wagner, and Wilkie (*2*) investigated the possibility of measuring the degree of deficiency in articulation. They state:

> If it were possible to devise a scale for measuring degrees of misarticulation, this scale would enable experimenters to compare the effectiveness of therapeutic procedures, and would enable clinicians to describe results of corrective procedures in more exact terms.

They considered several approaches to the problem of scale construction: (1) description of defective articulation in terms of the physical characteristics of the speech sound waves; (2) undertaking the phonetic analysis with the purpose of assigning weights to the various defects and thereby arriving at a score indicating a departure from acceptable speech; and (3) a rating scale using descriptive phrases to identify the various scale positions. They discarded the first because of its impracticability, the second because of its complexity, and the third because of its subjective and variable basis. They therefore developed a group of recordings with various degrees of misarticulation by means of the paired comparisons technique. A particular individual's speech was then judged in comparison with the devised set of records. The validity of the records used in the comparison is questioned here since they do not contain all the sounds in the English language. It is felt that the use of a rating scale for each sound with descriptive phrases to identify the scale positions need not be highly subjective and variable, that this was an unwarranted and untested assumption.

Templin (*11*), concerned about the method of testing articulation, recorded sounds as correct or incorrect, the latter consisting of indistinct, defective, substituted, and omitted sounds. Bangs (*1*) recorded speech responses as substitution, satisfactory, omission, and addition; Roe and Milisen (*8*), Sayler (*9*), Johnson (*6*) and Van Riper (*12*) have used a distortion recording.

Roe and Milisen (*8*) have indicated that a distorted sound may be considered part of a continuum in the development of a correctly articulated sound; that sounds are likely to be first omitted, then substituted, then distorted and finally correctly articulated. If this is true, it should be possible and clinically desirable to devise a descriptive scale of the degree of misarticulation as the sound approaches correctness. If such a scale were used a more adequate analysis of defective articulation might be obtained provided the scale was reliable.

Problem

It has been concluded that a finer measurement tool for defective articulation is needed than the customary recording of sounds as correct, distorted, substituted, or omitted. Milisen (7) advised the use of a seven point scale, beginning with a correctly articulated sound, then four levels of distortion, then substitution and omission. This scale recognized the previously neglected fact that a defective sound may vary in degree of distortion, as well as being substituted or omitted. The reliability of this method was not tested, however.

The problem, therefore, was to determine the reliability of evaluations made during articulation and stimulation testing according to the procedure originally recommended and used at Indiana University.

Materials and Subjects

Pictures for the articulation test were selected according to strict criteria. The goal of these criteria was to provide the examiners with material which would enable them to make their evaluations of articulation with a minimum of interference from extraneous variables.

The criteria used in selecting pictures which represented words containing the sounds to be tested were:

1. A vowel sound was immediately adjacent to the sound being tested. This eliminated the effect of one consonant sound on the production of another and the influence on evaluations of a consonant sound being tested by an adjacent consonant.

2. The tested sound occurred only once in any test word. This reduced confusions as to which sound was being evaluated.

3. The pictures represented words that were third grade reading level or lower (*10*) in order to assure simple testing material. This was also done to provide material for additional experimentation involving a comparison of printed and pictorial material in articulation testing.

4. Each picture was used in testing only one sound in one position. If a picture had been used to test two or more sounds, the first response to a picture could affect a later evaluation of a different sound tested by that same picture.

5. All pictures were of objects. This simplified the testing instructions. It was only necessary to ask what the picture was.

Sounds tested for reliability of evaluation were: [r l s f v k g tʃ] in the initial, medial, and final positions, as these were the only sounds that met all the criteria. Table 1 indicates the words used for the testing.

The individuals tested were 10 children whose articulation was defective and who had no organic disabilities. Seven male and three female children were used, with an age range from five years two months to eight years nine months and a mean age of six years seven months. All children were pupils in the School City of Bloomington, Indiana.

Three clinicians participated in this investigation of reliability. They were trained and had extensive experience in the methods of articulation testing used at the Indiana University Speech and Hearing Clinic.

Procedure

The initial step in the investigation of the reliability of a seven-point scale was agreement on the identification of each scale position which was as follows:

1. Sound was made correctly.
2. Sound was mildly indistinct and would probably be recognized as such by a speech teacher or someone else who was particularly interested in the English language, but would possibly not attract the attention of the average layman.
3. Sound was moderately indistinct and would perhaps attract the attention of the average layman to the speech of the individual, but still not be particularly annoying.
4. Sound was sufficiently indistinct so as to clutter the speech and to distract the average listener's attention from the speech content and occasionally cause him to misunderstand the words in which that sound occurred.
5. Sound was so severely indistinct that the average listener would not recognize it and would therefore fail to recognize many of the words in which the sound was used.
6. Sound was substituted for by another.
7. Sound was omitted.

During the testing, each of the examiners recorded a *one* for a correct sound; a *two* to a *five*, according to the degree of distortion, for a distorted sound; wrote in the phonetic symbol for a substituted sound; and recorded an *om* for an omitted sound.

There were three experimental conditions in this reliability investigation. Condition I was the administration of the test to each of the 10 individuals whose speech was defective. Conditions II and III consisted of listening to the tape recordings of Condition I.

During the first experimental condition, Condition I, one of the examiners served as the test administrator and sat across the table from the child to be tested. The other two examiners sat on each side of the test administrator. Each examiner recorded his evaluation of the speech response independently. The responses of each child who was tested were recorded on a portable Crestwood tape recorder with a Brush model BA-106 microphone.

TABLE 1. Words indicating the pictures used in the testing of initial, medial, and final positions of the sounds.

Sound	*Initial*	*Medial*	*Final*
r	rabbit	arrow	car
l	ladder	balloon	ball
s	saw	bicycle	house
f	fork	telephone	knife
v	valentine	seven	stove
k	cow	chicken	truck
g	gate	wagon	dog
tʃ	chair	matches	watch

The basic articulation section of the examination was administered first. In this section the child was told that he would be shown some pictures and he was to tell what they were: All positions of a sound were tested by means of pictures; the initial sound first, the medial second, and the final last. They were tested in the order in which they are given above, and in Table 1.

The stimulability portion of the articulation examination was then administered. Stimulation is defined as allowing the tested individual to see and hear the sound as it is produced by the examiner.

The child was stimulated three times with the [p] in isolation and asked to give one response. He was then stimulated three times with the nonsense syllable [ɑpɑ] and asked to give one response. This was done to familiarize the individual with the stimulability testing procedure, and his responses were not recorded.

If, at any time during the remaining

test administration, the tested individual gave more than one response to stimulation, only the first response was recorded.

The individual being tested was then stimulated, as above, first with the sound in isolation, then nonsense syllables, and then with the words representing the pictures used in the basic articulation section of the examination. The vowel [ɑ] was used in all nonsense syllables. This procedure was followed for each sound and the sounds were administered in the same order as in the basic articulation section of the examination. This testing procedure was followed for each of the 10 individuals.

In Condition II, the tape recordings of each child tested in Condition I were played back in random order 31 days after Condition I. The examiners evaluated the responses using the same procedures.

Condition III consisted of playing back the tape recordings of Condition I, randomly, 16 days after Condition II, which was 47 days after Condition I.

There was little probability that the three examiners remembered the speech responses of these 10 defective speakers since each examiner had to make 80 isolated evaluations for each child or a total of 800 evaluations in all. The possibility of an examiner's remembering the previously recorded responses was further guarded against by the lapse of time between the experimental conditions and the random playback of the tests.

Analysis of Data and Results

A rank was given to each of the evaluations of each sound in each position according to the seven-point scale previously agreed upon by the three examiners. These ranks were used in the statistical analysis of the data. Each evaluation was considered as one item; each test, therefore, consisted of 80 items. A total test score was obtained by adding the ranks attributed to each item for each test by each examiner. Thus, for each experimental condition each examiner had a total score for each of the 10 individuals tested.

The reliability of the examiners was investigated by comparing: (1) each examiner with himself between the experimental conditions; (2) each examiner with each other examiner within the experimental conditions;

TABLE 2. Consistency of each examiner's judgments with himself.

Experimental condition	*Examiner*	*rho*	*% exact agreement*	*% agreement ± one difference in rank score*
Condition I	D	.55	79	87
vs	H	.75	82	86
Condition II	K	.65	75	85
Condition II	D	.71	85	90
vs	H	.90	89	94
Condition III	K	.86	84	89
Condition I	D	.61	80	86
vs	H	.72	83	88
Condition III	K	.55	74	86

and (3) the examiners together within the experimental conditions.

First considered was the consistency of each examiner's judgments with himself. This was approached in two ways: (1) by the rank difference correlation coefficient (rho) of the total test scores, and (2) by a percentage of agreement on each item.

In considering the total test scores for each of the 10 individuals tested, a rank order was obtained for the 10 tested individuals for each examiner in experimental Conditions I, II, and III. How each examiner ranked the 10 tested individuals on the basis of their total test scores in one experimental condition was correlated with that same examiner's ranking of the tested individuals in the other experimental conditions. The results of the comparison of each examiner with himself may be found in Table 2.

Each item, for each tested individual in one experimental condition, was compared with the same item for the same tested individual in the other experimental conditions. Two percentages of agreement were obtained: (1) a percentage of exact agreement where the rank score on one item in one experimental condition was exactly the same on that same item in another experimental condition; (2) a percentage of agreement which allowed a difference of one rank in score to occur. Provision for one rank difference was made as it was believed that this is not a serious clinical error, especially in a seven-point scale, and it would permit the study of items evaluated in this manner. The consistency of each examiner's evaluations with exact agreement and with a plus or minus difference of one rank for the items in one experimental condition compared with another may also be found in Table 2.

When the consistency of each examiner with himself is considered, the results indicate that an examiner can be reliable with himself even when he is using the discriminative seven-point scale. The results also indicate that there is a greater reliability when tape recorded conditions are compared than when live and tape recorded situations are compared.

TABLE 3. Consistency of each examiner's judgments with every other examiner within the experimental conditions.

Experimental condition	*rho*	*% exact agreement*	*% agreement ± one difference in rank score*
Condition I			
D & H	.65	83	93
H & K	.75	77	88
K & D	.61	77	89
Condition II			
D & H	.82	82	88
H & K	.65	84	88
K & D	.67	81	87
Condition III			
D & H	.94	87	92
H & K	.84	81	88
K & D	.78	82	87

TABLE 4. Consistency of examiners' judgments among each other within the experimental conditions.

Experimental condition	rho_{av}	*W*	*% exact agreement*	*% agreement one difference in rank score*
Condition I	.67	.78	72	84
Condition II	.71	.81	76	81
Condition III	.85	.90	77	83

The second measure of the consistency of the examiners was the consistency of each examiner with each of the other examiners within each experimental condition. The measures of agreement were the rank difference correlation coefficient, the percentage of exact agreement, and the percentage of agreement plus or minus one difference in rank score. How two examiners ranked the 10 tested individuals on the basis of their total test scores was compared by means of the rank difference correlation coefficient. This is the same procedure followed for the consistency of each examiner with himself except each examiner was compared with every other examiner within the experimental conditions. The consistency of each examiner's judgments with every other examiner within the experimental conditions may be found in Table 3. These results indicate considerable consistency from one examiner to another.

Next examined was the consistency of the examiner's judgments when they were considered as a group and not separately. The consistency of judgment of each of the three examiner's ranking of the 10 tested individuals on the basis of their total test scores for one experimental condition was measured by the coefficient of concordance, or *W*.

Another measure of the rankings of each tested individual by the examiners as a group for each condition was the average rho which has a definite relation to *W*, the coefficient of concordance.

The percentage of agreement of the evaluations of the three examiners as a group was also determined. The percentage of complete agreement was determined for each item in each condition and in every position. The same was done for the evaluations which varied by not more than a plus or minus one rank.

These measures of consistency of the examiners' judgments among each

TABLE 5. Agreement of ranking determined by the total score given each individual by the examiners as a group.

Experimental condition	*rho*	rho_{av}	*W*
Condition I vs Condition II	.82		
Condition II vs Condition III	.95		
Condition I vs Condition III	.79		
Conditions I, II, and III		.85	.90

other within an experimental condition may be found in Table 4. These results indicate that all the examiners were highly consistent in their judgments when they were considered as a group.

In view of the previous result of greater reliability when tape recorded conditions were compared (Table 2), the experimental conditions were compared. This comparison was done by determining a ranking of the 10 tested individuals within an experimental condition by the examiners as a group. The three examiners' total scores for each tested individual within an experimental condition were added together and a new score determined for every tested individual. The tested individuals were then ranked on the basis of this new set of scores. This was done for each experimental condition. The ranks obtained for one experimental condition were then compared with the ranks obtained for the other experimental conditions. A rho_{av} and *W* were obtained in considering the three experimental conditions simultaneously. The results of this comparison of the experimental conditions may be found in Table 5.

These results indicate that the comparison of the tape recorded conditions have greater reliability than the comparison of the live and tape recorded conditions.

Conclusions

It was concluded that the evaluations made during articulation and stimulation testing according to the procedure originally recommended at Indiana University can be reliable.

Considerable agreement in evaluations of speech responses was observed from one condition to another for each examiner including the test administrator. This was demonstrated by reliability coefficients as well as percentage of agreement on each response.

Considerable agreement in evaluations of speech responses was observed from one examiner to another for each condition. This was demonstrated by reliability coefficients as well as percentage of agreements.

Greater agreement was observed in evaluations of speech responses of the two tape recorded conditions than in the live testing condition. This was shown by correlations and percentage of agreements. This was true when each examiner was compared with himself as well as when he was compared with the other examiners. It was concluded that the greater variation in the evaluations made of live speech than tape recorded speech was due to the presence of more variables in the live speech situation. Nevertheless the relationship between the live and recorded speech was high enough to warrant the use of tape recording as a medium in training students who are learning to evaluate articulation responses.

References

1. Bangs, J. L. A clinical analysis of the articulatory defects of the feebleminded. *JSD*, 7, 1942, 343-356.
2. Curry, R., L. Kennedy, L. Wagner and W. Wilkie. Phonographic scale for the measurement of defective articulation. *JSD*, 8, 1943, 123-126.
3. Henderson, F. M. Accuracy in testing the articulation of speech sounds. *J. educ. Res.*, 31, 1938, 248-356.
4. Irwin, O. C. and H. P. Chen. A reliability study of speech sounds observed in the crying of newborn infants. *Child Developm.*, 12, 1941, 351-368.
5. Irwin, O. C. and T. Curry. Vowel elements in the crying vocalization of infants under ten days of age. *Child Developm.*, 12, 1941, 99-109.

6. JOHNSON, W., S. BROWN, J. CURTIS, C. EDNEY AND J. KEASTER. *Speech Handicapped School Children.* New York: Harper, 1948.
7. MILISEN, R. Principles and methods of articulation testing. *Speech and Hearing Therapist,* Feb., 1945, 6-10.
8. ROE, V. AND R. MILISEN. The effect of maturation upon defective articulation in elementary grades. *JSD,* 7, 1942, 37-50.
9. SAYLER, H. The effect of maturation upon defective articulation in grades seven through twelve. *JSHD,* 14, 1949, 202-207.
10. STONE, C. *Stone's Graded Vocabulary for Primary Reading.* St. Louis: Webster Publ., 1941.
11. TEMPLIN, M. Spontaneous versus imitated verbalization in testing articulation in preschool children. *JSD,* 12, 1947, 293-300.
12. VAN RIPER, C. *Speech Correction, Principles and Methods.* New York: Prentice-Hall, 1939.

III

THE INFLUENNCE OF ORAL VERSUS PICTORIAL PRESENTATION UPON ARTICULATION TESTING RESULTS

Katherine Snow and Robert Milisen

AN IMPORTANT problem in articulation testing involves the technique to be used in eliciting the particular sound to be tested.

A test can be conducted by holding a conversation with the child and noting the sounds as they appear in words spontaneously uttered. Though this is a good test of the actual propositional speech, it is inefficient and time-consuming as it may take a tremendous amount of conversation to elicit some of the less common sounds. For older children, articulation can be readily tested by having the child read aloud words or sentences containing the sounds to be tested.

For younger children, the most common method consists in showing the child pictures which illustrate words within the child's vocabulary, which words contain the sounds to be tested. Since some words are difficult to illustrate, many investigators (*1*, *5*, *9*) ask the child to repeat these words after oral stimulation. The question then arises as to whether the child will tend to correct his errors after the stimulation. Morrison (*4*) and Templin (*8*) offer some experimental evidence that the articulation is little if any different when an oral pattern has been given, than it is when it is elicited pictorially.

Practical clinical experience leads some speech clinicians to believe, however, that the oral stimulation will change the pattern of response (*3*). Since oral stimulation is often effective in therapeutic situations it seems likely to have some effect in testing situations. The purpose of this study was to determine whether defective speaking children give the same oral responses to words presented orally as to words presented pictorially or in writing.

Subjects

There were 164 subjects, consisting of all of the children referred to the examiner by their teachers as being defective in articulation from the first, second, seventh and eighth grades from all of the white schools of the Fayette County, Kentucky, public schools.

The patrons of the schools from which the subjects were drawn ranged from poor tenant farmers to prosperous suburbanites, but the average socio-economic status of the group was well above national averages. None of the children were receiving or had ever received any speech therapy.

Table 1 represents a summary of the grade level and sex of the subjects.

Procedure

The 164 children were tested on 25 consonant sounds and blends, the same sounds used by Roe and Milisen (*6*) and Sayler (7). The first and second grade children were given a picture articulation test and a test of imitated verbalization. The seventh and eighth grade children were asked to read single words which had been typed on cards and also tested with imitated verbalization.

For the purposes of this study, the articulation test using imitated verbalization will be called the oral test and the articulation test using pictures to elicit the sounds will be called the picture test. Since, in the seventh and eighth grades, the subjects read words instead of responding to pictures, this reading test will be considered along

TABLE 1. Grade level and sex of subjects.

Grade	*Boys*	*Girls*	*Total*	*%**
I	41	17	58	10.6
II	35	21	56	8.1
VII, VIII	36	14	50	5.3
Total	112	52	164	

*Percentage of the total enrollment in the grade.

with the picture test in making an analysis of the results.

The method used in administering the picture test was similar to that used by Roe and Milisen (*6*). The reading articulation test for the seventh and eighth grades differed from that used by Sayler (7) in that single words, instead of sentences, were used. The oral test was patterned after the one used by Templin (*8*), but the examiner simply said: 'I am going to say some words and want you to say them after me,' whereas Templin asked to child to ' . . . say it after me so I'm sure to hear it just right.'

The children were given two complete tests, one day apart. Test I was half oral and half picture, while in Test II, the test was reversed; that is, the sounds tested by the use of pictures in Test I were tested by the oral method in Test II and vice versa. An effort was made to divide the words on the tests in such a way that voiced and unvoiced sounds were well distributed and similar sounds were scattered. In all testing sessions alternate children were given Test I and Test II, and each child was then given the other test on the following day. The order of presentation of the picture and oral sections of the tests was also alternated. In this way each child was tested by both methods for each sound, and it was hoped that any practice effect was cancelled out.

Since each test was recorded on a separate blank, opportunity for the examiner to be influenced by a previous judgment was minimized. No determination of the reliability of the examiner was made at the time the data were being gathered, but the examiner later participated in the study by Wright, above, in which the reliability of the three examiners with themselves and as a group was shown to be high.

Only one sound was tested in each word. The same words were used for each sound in all tests, except in a few cases where alternates were available. Some words do not elicit oral responses from pictures as readily as might be desired. Alternates were available for such words and for others reported by Roe and Milisen (*6*) as possibly unsatisfactory test words. They were used in this study in the few cases where the child did not respond to the first test word. In some instances a child might not respond at all to the pictorial stimulation for a sound, but in no case was the word spoken for the child during the picture or reading tests. In the oral test the child was not asked to look at the examiner or to listen carefully, but only to repeat each word after it was said once.

Testing was usually done in a relatively quiet room with only the child and the examiner present. No attention was called to the fact that a speech test was being given, and only a few of the seventh and eighth graders seemed to realize that it was a speech test.

The responses were recorded on the test blanks as omission, substitution (the substituted sound indicated phonetically), or degree of distinctness. For the latter a scale of one to three was used, one being a correct sound, two a moderately indistinct sound, and three a sound so indistinct that the average listener might not recognize it.

Results

The test responses for the sounds [s z θ ð r l tʃ dʒ k g f v p b t d ʃ ŋ] in all positions were tabulated in the following manner: A omission was recorded as (5), a substitution as (4), a very indistinct sound as (3), a moderately indistinct sound as (2), and a correct sound as (1). Clinical observation indicates that the direction of progress toward correction of a defective sound is from omission to substitution to indistinctness, though, of course, a sound does not necessarily go through each step. Roe and Milisen (*6*) came to this conclusion, feeling that their results gave a clue to the development of correct sounds from incorrect ones, when no specific training was given. In another section of the present study, it was found that in a six-month period without speech therapy, omissions changed to substitutions in first grade children five and one half times, and in second grade children seven and one half times more frequently than substitutions changed to omissions. Substitutions changed to indistinct sounds in first grade children nine times more often and in second grade children 23 times more often than substitutions changed to omissions.

While this does not show that a sound necessarily follows a particular order, it indicates a general trend from omission to substitution to indistinctness, and justifies the use of the rank order scale as described above.

These data can be analyzed in many ways. The method used by earlier investigators compared correct with incorrect responses, no attempt being made to subdivide the incorrect responses. This procedure has limitations in that many facts pertinent to diagnosis of defective articulation may be found in different types of articulation errors. Two methods of scoring were used in this study. In Table 2, the older method of comparing correct versus incorrect responses was used. This was the method used in Templin's study (*8*). In Table 3, account was taken of four degrees of

TABLE 2. Differential responses between initial oral and initial picture tests based on correct versus incorrect responses.

	Better Oral		*Better Picture*		*No Change*		*Total*	
	Number	*Per Cent*	*Number*	*Per Cent*	*Number*	*Per Cent*	*Number*	*Per Cent*
Grade 1	393	31	144	11	746	58	1283	100
Grade 2	265	28	114	12	565	60	944	100
Grades 7, 8	195	50	82	21	111	29	388	100
Total	853	33	340	13	1422	54	2615	100

TABLE 3. Differential responses between initial oral and initial picture tests based on better, but not necessarily correct responses.

	Better Oral		Better Picture		No Change		Total	
	Number	Per Cent	Number	Per Cent	Number	Per Cent	Number	Per Cent
All Sound Positions								
Grade 1	537	42	280	22	466	36	1283	100
Grade 2	371	39	195	21	378	40	944	100
Grades 7, 8	223	57	95	24	70	18	388	99
Total	1131	43	570	22	914	35	2615	100
Initial Sounds								
Grade 1	147	45	50	15	131	40	328	100
Grade 2	100	39	39	15	120	46	259	100
Grades 7, 8	50	67	13	17	12	16	75	100
Total	297	45	102	15	263	40	662	100
Medial Sounds								
Grade 1	185	42	101	23	153	35	439	100
Grade 2,	115	39	65	22	118	40	298	101
Grades 7, 8	70	60	25	22	21	18	116	100
Total	370	43	191	22	292	34	853	99
Final Sounds								
Grade 1	205	40	129	25	182	35	516	100
Grade 2	156	40	91	24	140	36	387	100
Grades 7, 8	103	52	57	29	37	19	197	100
Total	464	42	277	25	359	33	1100	100
Visible [f v θ ð p b]								
Grade 1	162	37	99	23	176	40	437	100
Grade 2	129	39	56	17	144	44	329	100
Grade 7, 8	71	55	32	25	25	20	128	100
Total	362	40	187	21	345	39	894	100
Partially Visible [t d k g ŋ tʃ dʒ r l]								
Grade 1	271	47	122	21	186	32	579	100
Grade 2	174	42	94	22	150	36	418	100
Grades 7, 8	85	57	39	26	26	17	150	100
Total	530	46	255	22	362	32	1147	100
Non-visible [s z ʃ]								
Grade 1	104	39	59	22	104	39	267	100
Grade 2	68	35	45	23	84	43	197	101
Grades 7, 8	67	61	24	22	19	17	110	100
Total	239	42	128	22	207	36	574	100

articulation errors, in that, for example, a change from an omission (5) to a substitution (4), or from a substitution (4) to a moderately indistinct sound (2), was considered a better response even though it was incorrect in both instances.

Table 2 presents the number and percentage of responses evaluated simply as correct or incorrect (the degree of incorrectness was not recorded). The sounds which were 'Better Orally' were those which were correct on the oral test and incorrect on the picture test. The sounds which were 'Better Pictorially' were those which were correct on the picture test and incorrect on the oral test. The sounds which were placed in the 'No Change' column were sounds which were incorrect on both tests. Instances where responses to sounds were correct in both tests were not included in either Table 2 or Table 3.

Table 3 presents the number and percentage of responses which were better according to the previously described rank order scale, and also the percentage of responses which were equally incorrect on both tests according to the scale. The 'Better Oral' responses were those better on the oral than on the picture test. The 'Better Picture' were the reverse. The 'No Change' responses were equally defective on both tests. These percentages are shown for all sounds, and for the sounds in the initial, medial and final positions, as well as for the visibility of the focal articulation points.

The similarity in the trends shown in Tables 2 and 3 is quite striking and would seem to highlight the very definite trend toward better articulatory responses after oral stimulation as compared with pictorial stimulation. The 'No Change' column in Table 2 shows much higher percentages than the corresponding column in Table 3 because, for example, a change from an omission (5) on one test to a moderately indistinct sound (2) on the other would appear in this column since the responses were incorrect in both tests. In Table 3 any change on the rank order scale would put the response into either the 'Better Oral' column or the 'Better Picture' column. Only identical incorrect responses appear in the 'No change' column in Table 3.

Where there was a difference in the response to the oral and picture tests, as shown in Tables 2 and 3, the better response was, roughly, twice as likely to occur on the oral test. The older children, in grades seven and eight, showed a much greater differential in favor of the oral test than did the younger children. Whether this reflects a greater responsiveness to oral stimulation on the part of the older children or is somehow connected with the fact that they read words instead of responding to pictures, or has some other explanation, would be impossible to say on the basis of the available data.

From a comparison of the positions of the sounds in the words, there appears to be a very slight trend for the initial sounds to be most influenced by oral stimulation, then medial and final sounds, in that order. The visibility of the focal articulation points of the sounds seems to have no particular bearing on the production of the sounds after oral stimulation, but it must be kept in mind that the child was not asked to look at the examiner. Some children looked and listened, some only listened, to the one production of the word.

An articulation score was figured for each child by adding the rank order values assigned to his responses for each sound in each position and dividing by the number of responses. Thus the score is the mean of the values of the rest responses, and can conceivably vary between 1.00 (all responses correct) and 5.00 (all tested consonants omitted). This actually involved the use of the rank order scale of accuracy of articulation, as a numerical scale. There is no simple way to construct an adequate, weighted numerical rating scale for degrees of accuracy of articulation. Such factors as frequency of occurrence of each sound in each position,

TABLE 4. Mean articulation scores for picture and oral tests.

Grade	N	Picture Test	Oral Test	Diff.	S.E.	t
I	58	2.04	1.85	.19	.0326	5.84
II	56	1.80	1.68	.12	.0224	5.35
VII, VIII	50	1.24	1.15	.09	.0237	3.79

the relation of a sound to an adjacent sound, the nature of each misarticulation and its relationship to the development of speech and its effect upon the listeners, and many others, would need evaluation. It was felt that unless a complete separate study could be undertaken to establish a numerical scale (and limitations of time made this impossible), it was better to use the simplest method as the one least likely to introduce unnecessary error. Therefore, the rank order scale described previously was used also as a numerical scale in arriving at the articulation score.

Table 4 shows the mean articulation scores for the picture and oral tests for Grade I, Grade II, and Grades VII and VIII combined, with the differences between the picture and oral test scores. As would be expected, the scores decrease (improve) with increase in grade. There is also a rather consistent difference between the picture and oral test scores, which difference also decreases with increase in grade level.

The *t*-test[1] for the significance of the difference in the means of related pairs (in this case the pairs consist of the same child given two types of test) indicates a high degree of confidence that the difference is not due entirely to chance. It is significant at less than the one per cent level for all grades tested, but is most significant in the lower grades.

[1]
$$t = \frac{M_O - M_H}{\sqrt{\frac{\Sigma d^2}{n(n-1)}}}$$

Where M_O is the observed mean, M_H is the hypothetical mean, Σd^2 is the summation of the squares of the deviation from the mean of the differences between the pairs, and n is the number of pairs.

Summary and Conclusions

In this study, 114 defective speaking school children from Grades I and II were each given an oral and a picture articulation test, and 50 defective speaking school children from Grades VII and VIII were each given an oral and a reading articulation test.

The children did not give the same oral responses to the two types of test. There was a consistent differential in favor of better responses to the oral test. This differential was present at all grade levels tested and seems to indicate that the small amount of stimulation present in the oral test was influencing the articulation responses by aiding the children to form their sounds more correctly.

The results of this study indicate that of the two tests the picture, not the oral, test should be preferred when testing the articulation of children.

References

1. HENDERSON, F. A study of the articulation of consonants by normal institu-

tionalized children. Ph. D. Dissertation, Univ. Wis., 1935.
2. LINDQUIST, E. F. *Statistical Analysis in Educational Research.* Boston: Houghton Mifflin, 1940.
3. MILISEN, R. Principles and methods of articulation testing. *Speech and Hearing Therapist*, Feb., 1945, 6-10.
4. MORRISON, C. E. Speech defects in the speech of young children. *Psychol. Clin.*, 8, 1914, 138-142.
5. POOLE, I. The genetic development of articulation of consonant sounds in children's speech. Ph.D. Dissertation, Univ. of Mich., 1934.
6. ROE, V. AND R. MILISEN. The effect of maturation upon defective articulation in the elementary grades. *JSD*, 7, 1942, 37-50.
7. SAYLER, H. K. The effect of maturation upon defective articulation in grades seven through twelve. *JSHD*, 14, 1949, 202-207.
8. TEMPLIN, M. Spontaneous versus imitated verbalization in testing articulation in preschool children. *JSD*, 12, 1947, 293-300.
9. TEMPLIN, M. AND M. STEER. Studies of growth of speech of preschool children. *JSD*, 4, 1939, 70-79.

IV

THE EFFECT OF VISUAL, AUDITORY AND COMBINED VISUAL-AUDITORY STIMULATION UPON THE SPEECH RESPONSES OF DEFECTIVE SPEAKING CHILDREN

Davis A. Scott and Robert Milisen

ALTHOUGH speech therapists are almost in complete agreement concerning the primary goals of therapy for children with defective articulation, by no means are they agreed upon the procedures that should be utilized in attaining these goals. A large number of methods have been presented in the literature and are advocated by various centers of speech correction. In addition, each therapist creates his own variations. Basically, however, all techniques utilized may be classified under three broad, general methods: The phonetic-placement method, the moto-kinesthetic method, and the stimulus or stimulation method.

The underlying principle of the phonetic-placement method involves having the student attend to the movements and positioning of the articulatory structures; the goal is conscious control of movement and positioning. In general, the procedures include: the use of visual or verbal instructions, the use of exercises for the speech musculature, and the modification of other sounds already in the child's repertoire.

Whereas phonetic-placement involves merely instructions and exercises in order to teach a static position, the moto-kinesthetic method involves the actual manipulation of the student's articulatory mechanism, either mechanically or manually. The procedures include the use of instruments, applicators, or the fingers of the therapist in order to attain the proper position of the mechanisms.

The stimulation method, although not incorrectly named, does mislead the person when he considers other methods. For, in a strict psychological sense, any technique or device or procedure that is used to teach a child how to say a sound may be considered as a stimulus. In the stimulation method, however, the stimuli presented to the student are auditory and visual: a model sound is presented by the therapist and the student imitates what he sees and hears.

It is the purpose of this study to obtain some experimental evidence which would point to the effect of visual and auditory stimulation upon the speech responses of defective speaking children. No attempt will be made to draw comparisons between the stimulation method and the other methods of therapy. The attempt will be to secure information in order to evaluate the worth of the stimulation method in teaching new sounds. More specifically, the question to be answered is: How effective are visual stimuli alone and auditory stimuli alone in comparison with combined visual and auditory stimulation?

Procedure

Subjects. The subjects used in this experiment were 64 of the defective speaking children examined by the Indiana University Traveling Speech and Hearing Clinic in the summer of 1951. Only children with defective articulation were accepted as subjects. Children were excluded from the experimental group if they had one or more of the following problems: a vision defect, a hearing problem, cerebral palsy, cleft palate, or stuttering.

Furthermore, the sample of subjects was limited to include only those children whose faulty articulation was mainly a learning problem. This limitation excluded those children who had personality or emotional problems in addition to, or because of, their speech handicap. It also served to exclude those children who because of lack of cooperation or lack of attention could not complete the examinations required.

It was felt that the group of subjects selected constituted a normal sample of the children with defective articulation, whose problems were mainly educational in nature and were not complicated by other handicapping conditions. The ages of the children ranged from four to 14 with over half of the subjects either seven or eight years of age. The range in grade placement extended from kindergarten to sixth grade with over half of the subjects in either the first or second grade. Five children had not entered school. In terms of mental maturity, six of the subjects were classified as having borderline intelligence, 11 as having below average intelligence, 39 as having average intelligence, six as having above average intelligence, and two as having superior intelligence. Of the 64 subjects, 42 were boys and 22 were girls.

Selection of Sounds. For the purpose of this experiment the examiner was interested only in the testing and stimulating of the following sounds: [r l s z k g f v]. These sounds were selected for further study because they contain most of the variable characteristics of consonant sounds. The important factors they contained as far as this study is concerned are the degree of visibility and audibility since these factors facilitate the limitation of sounds. The focal articulation points of the [f] and [v] sounds are readily visible; the focal articulation points of the [k g r l] sounds can be made more visible but under ordinary circumstances are not readily visible; the focal articulation points of the [s] and [z] sounds are non-visible. Other variables these sounds contain are that some are voiced while others are unvoiced; some are plosives, some are fricatives, some are sibilants, and some are semivowels.

Tests. A picture articulation test was administered to determine which sounds were defective. The picture test was constructed in the form of a notebook; each page was devoted to the examination of one sound as it appeared in the three positions of words.

A stimulation test was administered next to determine how well the subject could produce each of his defective sounds in isolation. The subject was first asked to produce the sound following combined visual and auditory stimulation; that is, the experimenter said the sound and the child *looked*, *listened*, and *imitated*. Next, the child imitated the sound following a visual stimulus alone. Last, he produced the sound following an auditory stimulus alone.

The stimulation test was begun on the [θ] sound so as to give the subject an indication of what was required of him and to serve as a practice session. The results of the [θ] stimulation were not included as part of the experimental data. If the [θ] sound was not defective, the subject would be given practice on another of the sounds being studied; the results of such stimulation were not then included in the data.

Directions. For the picture articulation test, the examiner said:

> I am going to show you a book full of pictures. I am going to point to some of them and I want you to tell me what they are.

In some instances it was necessary to ask leading questions in order to obtain the desired word-responses; however, at no time was the desired word said by the examiner before the subject said it.

For those conditions on the stimulation test in which combined visual and auditory stimuli were presented, the examiner said:

> I want you to watch my mouth very closely and to listen very carefully. I am going to say some sounds and I want to see how well you can say what I do.

For those conditions in which a visual stimulus alone was presented, the examiner said:

> Now watch my mouth very closely. I am going to whisper a sound and I want you to say the same sound out loud after me.

The sounds were not actually whispered; the focal articulation points were formed without any of the auditory characteristics.

For those conditions in which an auditory stimulus alone was presented, the examiner said:

> Now listen very carefully. I am going to hide my mouth so that you cannot see what I say. I want to see if you can say the same sound after me just like I do.

The examiner's mouth was hidden by means of holding a 3 × 5 white index card in front of his lips.

After giving the directions for one sound, it was often unnecessary to repeat them; however, if they were needed again, they were repeated.

Method of Recording Responses. The responses to the basic articulation tests and to the stimulation tests were recorded in the same manner. If the subject omitted a certain sound in a certain situation, a symbol denoting omission (O) was recorded on the test form. If the subject substituted one sound for another, the phonetic symbol for the sound substituted was recorded. If the subject distorted a sound, a symbol denoting the degree of distortion (5, 4, 3, or 2) was recorded. If the subject produced a correct sound, (1) was recorded.

Reference to this examiner's reliability in judging and recording responses may be found in the preceding article by Wright concerning the reliability of examiners' judgments of speech responses.

Results

Since it was impossible to set up any controls to regulate the order of presentation of the stimuli, that is, since the visual-auditory stimulation always came first, then visual stimuli, and then auditory stimuli, it is impossible to apply any statistical analysis to the data. Therefore, the experimenter tabulated the frequency of various types of response to the different types of stimulation and displayed these results in a tabular form for each sound. Although this method of analysis does not allow for the determination of significance, the tabular presentation does summarize the results well and points to trends that existed in the data.

The responses to stimulation of the sound in isolation were first classified as to whether they were correct responses, distortions, substitutions, or omissions. Then the frequency of these types of response following each type of stimulation was tabulated for each sound. The results of these tabulations appear in Table 1. It should be noted that of the 64 subjects, 33 received stimulation on the [f] sound, 35 on the [v], 37 on the [k], 32 on the [g], 27 on the [r], 40 on the [l], 45 on the [s], and 35 on the [z].

The results for the three types of stimulation of [f] are practically identical. Taking the practice effect into consideration, one might expect the

TABLE 1. Frequency of various types of response following visual-auditory, visual, and auditory stimulation of the sounds tested.

Types of Response	*Visual-Auditory*	*Visual*	*Auditory*	*Visual-Auditory*	*Visual*	*Auditory*
		f			**v**	
Correct responses	33	32	30	33	23	29
Distortions	0	0	1	0	2	3
Substitutions	0	0	2	2	10	3
Omissions	0	1	0	0	0	0
		k			**g**	
Correct responses	35	28	30	30	21	27
Distortions	0	3	2	0	2	2
Substitutions	2	6	5	2	9	3
Omissions	0	0	0	0	0	0
		r			**l**	
Correct responses	13	6	11	31	22	15
Distortions	14	15	13	7	9	12
Substiitutions	0	5	3	2	8	13
Omissions	0	1	0	0	1	0
		s			**z**	
Correct responses	36	23	31	25	11	21
Distortions	9	15	9	6	9	10
Substitutions	0	6	5	4	13	4
Omissions	0	1	0	0	2	0

results for either the visual or auditory stimulation to show more correct responses than the visual-auditory results; however, this is not the case. Therefore, we may assume that the combined visual-auditory stimulation is more effective in eliciting correct responses for the [f] sound since the results were better and there was no practice effect. There is no indication of which factor, the visual or auditory, is more helpful in facilitating the imitation of this sound or which isolated stimulus presents the greater cue to the subject.

The results for [v] indicate that more correct and fewer incorrect responses were produced following visual-auditory stimulation. Although the table indicates that the auditory stimulus produced the second best results, it should be pointed out that all of the 10 substitutions recorded for the visual stimulus were substitutions of the voiceless cognate, [f] for [v]. Considering this fact, it appears that neither the visual nor auditory stimulus is more effective in producing the correct focal articulation point; however, the auditory stimulus often appears necessary in the production of a sound that is auditorily as well as visually correct.

The results for [k] indicate that more correct responses and fewer incorrect responses were produced following visual-auditory stimulation. Fewer correct and more incorrect responses were made following the visual stimulus; however, there appears to be no definite indication of which factor, visibility or audibility, is the more effective in facilitating the imitation of the sound.

The results for [g] indicate that more correct responses and fewer incorrect responses were produced following visual-auditory stimulation. Fewer correct responses and more incorrect responses were produced following visual stimulation alone. Once again, however, a number of these incorrect responses were substitutions of the voiceless cognate. Therefore, it appears that the visual stimulus enables the subject to produce a sound that is visually correct although auditorily incorrect. No further clue was present that would indicate the relative effectiveness of visual versus auditory stimulation.

The results for [r] indicate that more correct and fewer incorrect responses were produced following visual and auditory stimulation. The responses that occurred following auditory stimulation were almost as good. Fewer correct responses and more incorrect responses were produced following visual stimulation alone. The results indicate that the auditory factor may be more important than the visual factor in facilitating the imitation and learning of the [r] sound.

The results for [l] indicate that far more correct sounds and fewer incorrect sounds were produced following visual-auditory stimulation. Fewer correct responses and more incorrect responses were produced following auditory stimulation. These results indicate that visual stimulation may be more effective than auditory in enabling a subject to produce a correct [l] sound.

More correct sounds and fewer incorrect sounds for [s] were produced following visual-auditory stimulation. The responses were nearly the same following the auditory stimulation. Far fewer correct responses and more incorrect responses followed the visual stimulation. These results indicate that the auditory stimulus may be more important than the visual stimulus in enabling a subject to produce a correct [s] sound.

More correct responses and fewer incorrect responses for [z] were produced following visual-auditory stimulation. Far fewer correct sounds and more incorrect sounds occurred following visual stimulation. These results indicate that the auditory stimulus may be more important than the visual factor in the production of the [z] sound.

Summary

This analysis of the effect of various types of stimulation indicates that the combined visual-auditory stimulation is the most effective. For all sounds, more correct responses and fewer incorrect responses were produced following the visual-auditory stimulation. The auditory stimulus elicited more correct responses and fewer incorrect responses than did the visual stimulus for sounds with relatively non-visible focal articulation points, the [s] and [z] sounds, and for one sound with a focal articulation point that can be made more visible, the [r] sound. The visual stimulus elicited more correct responses and fewer incorrect responses than did the auditory stimulus for one sound with a focal articulation point that can be made more visible, the [l] sound. The two types of stimulation elicited approximately the same responses for

the other two sounds with focal articulation points that can be made more visible, the [k] and [g] sounds, and for those sounds with relatively visible focal articulation points, the [f] and [v] sounds. Thus, the visual stimulus appears to be just as important as the auditory stimulus. For the voiced sounds [v] and [g], however, the auditory stimulus appears slightly more effective in eliciting responses that are auditorily correct; the visual stimulus alone elicits a number of substitutions of the voiceless cognate.

V

SPONTANEOUS IMPPROVEMENT IN ARTICULATION AS RELATED TO DIFFERENTIAL RESPONSES TO ORAL AND PICTURE ARTICULATION TESTS

Katherine Snow and Robert Milisen

DIFFERENCES which have been shown to exist between the oral responses given by defective speaking children to oral and picture articulation tests, as described by Snow and Milisen, above, raise questions as to the relationships between these differences and prognosis and therapy in articulation disorders.

Successful articulation therapy, no matter what methods or techniques are used, involves the production by the child of a more adequate articulatory response. The same is true of improvement in articulation that takes place without formal speech therapy. In articulation testing, minimum assistance in making a sound is given to the child when a word is elicited by use of pictures. More than minimum assistance in making a sound is given to the child when a word is elicited by saying it first for the child, since he has a correct oral pattern and perhaps a correct visual pattern to follow. Since the differential to the responses to the two types of tests would appear to be the result of the assistance given by the presentation of the correct pattern, there would seem to be a possibility that the child who gives a better response to the oral pattern would be able to make more progress in correcting his articulation than a child who gives the same incorrect response after oral stimulation that he gives to a picture. There would also seem to be a possibility that those sounds that were made better after the limited oral stimulation would be more susceptible to improvement.

The purpose of the present study is to investigate, in defective speaking children, the degree of spontaneous improvement in articulation as it is related to the differential responses to oral and pictorial articulation tests. This comparison was made for each child and for each sound.

Procedure

The subjects consisted of 81 defective speaking first and second grade children from the Fayette County, Kentucky, public schools, who had not received speech therapy. These are the children described in the Snow and Milisen study, above, who were available for retesting after a six-month interval.

The 81 children were tested on 25 consonant sounds and blends, the same sounds used by Roe and Milisen (2). In the fall each child was given a picture articulation test (called for the purposes of this study the Initial Picture Test) and a test of imitated verbalization (called for the purposes of this study the Initial Oral Test). Six months later, in the spring, each child was retested with the same tests. Of the spring tests only the results for the Final Picture Test were used for this study.

Results

The test responses for the sounds [s z θ ð r l tʃ dʒ k g f v p b t d ʃ ŋ] in all positions were tabulated and values for each type of response were assigned in the following manner: An omission a value of (5), a substitution a value of (4), a very indistinct sound a value of (3), a moderately indistinct sound a value of (2), and a correct sound a value of (1). A justification for use of this scale of values may be found in the earlier paper.

Four kinds of distributions involving differential scores were obtained by comparing results between tests.

The first type of distribution was obtained by comparing the results each *child* made on the Initial Picture Test with those he made on the Initial Oral Test. The second type of distribution was obtained by comparing the results each *child* made on the Initial Picture Test with the results he made six months later on the Final Picture Test. The third type of distribution was obtained for all sounds by comparing the results for each *sound* in each position on the Initial Picture Test with those for the Initial Oral Test. The fourth type of distribution was obtained for all sounds by comparing the results for each *sound* in each position on the Initial Picture Test with those made six months later on the Final Picture Test.

Differentials between results on the two Initial Tests represent the amount of difference obtained between oral and pictorial stimulation. This result helps to answer the question, 'Which kind of test gives a child the most assistance in articulating?'

Differentials between results on the Initial and Final Picture Tests represent the spontaneous improvement in a child's articulation during the six-month period between the administration of the Initial and the Final Picture Tests.

The first two types of distributions of differential scores were computed in the following manner: An articulation score for each child on each test was obtained by assigning, to each response he made to each test word, a value. These values, previously described, were added and the sum was divided by the number of responses. The mean of these values represented his articulation score and could range from a 1.00 for a perfect articulation score to a 5.00 for a score containing all omissions.

TABLE 1. Correlations between the differences in the means of the children's articulation scores on the Initial Picture and Oral Tests and the differences in the Initial and Final Picture Tests.

Grade	*N*	*r*
I	38	.71
II	43	.53

The first distribution of differential scores was obtained by subtracting the articulation score made by each child on the Initial Picture Test from that made by him on the Initial Oral Test.

The second distribution of differential scores was obtained by subtracting the articulation score made by each child on the Initial Picture Test from that made by him on the Final Picture Test. Correlations[1] were calculated between these two distributions of differential scores for each grade. As shown in Table 1 for Grade I, this correlation was .71 and for Grade II it was .53. According to Lindquist (*1*), these correlations are significant at the one per cent level of confidence.

In general, children made more errors, or more severe errors on the Initial Picture Test than on the Initial Oral Test. They received oral stimulation from the oral test, but none from the picture test. For the most part, those who had the greatest differential score between the two tests were the ones who showed the greatest spontaneous improvement in artic-

[1]

$$r_{xy} = \frac{\bar{x}\,\bar{y}}{n\sigma x \sigma y},$$

where $\bar{x}$ and $\bar{y}$ are the deviations from the respective means, and n is the number of cases ulation as measured six months later by the Final Picture Test.

This would indicate a probability that the difference in a child's responses to an oral and a picture articulation test could be used as one valuable factor in predicting his progress in correcting his articulation errors.

Another question which arises is whether sounds and their positions in words affect the prediction of spontaneous improvement. To study this question, the third and fourth types of distributions of differential scores were obtained for all sounds, in each of the three positions.

The third type of distribution of differential scores was computed in the following manner: A mean articulation score for each sound in each position was obtained for the Initial Picture Test. Each score was calculated by adding the values of all responses to a sound in a given position and dividing by the number of responses. This same procedure was followed for the results on the Initial Oral Test. A difference score was obtained by subtracting the mean for a sound in a given position on the Initial Picture Test from the mean on the Initial Oral Test. A distribution of these differential scores was made.

The fourth type of distribution of differential scores was obtained in a similar manner, except that the Initial Picture Test results were compared with the Final Picture Test.

The differential scores for the third type of distribution were correlated with those obtained for the fourth type of distribution for each position of sounds in words for each grade. As shown in Table 2, these correlations indicate that, in general, sounds produced better in the Initial Oral Test than in the Initial Picture Test also showed the greatest spontaneous improvement as measured six months later in the Final Picture Test.

TABLE 2. Correlations between differences in means of Initial Picture and Oral Tests, and Initial and Final Picture Tests, by individual sounds.

Positions of Sounds	*Number of Sounds*	*r* *Grade I*	*Grade II*
Initial	17	.67	.77
Medial	18	.44	.47
Final	18	.70	.52

The significance of the correlations presented in Table 2 are as follows: The initial positions of sounds in both Grades I and II were significant at the one per cent level. The medial sounds for both grades were significant at the five per cent level. The final sounds for Grade I were significant at the one per cent level and for Grade II at the three per cent level.

Summary and Conclusions

In this study, 81 defective speaking school children from Grades I and II were each given an oral and a picture articulation test. They were retested with the same tests after six months, during which time they received no speech therapy.

The study indicates a probability that the difference in a child's responses to an oral and a picture articulation test could be used as a valuable factor in predicting his progress in correcting his articulation errors. It also indicates that, to a considerable extent, sounds which are produced better in an oral than in a picture articulation test, are the ones which will show the most spontaneous improvement in articulation.

In assessing the results of this study it must be kept in mind that the differential responses to the oral and picture articulation tests were produced after a minimum of oral stimulation. The child was asked to repeat a word after it was said only once. He was not asked to look at the examiner. If this very limited stimulation is sufficient to show the predictive value that is here indicated, it would seem that a more carefully designed oral and visual stimulation test, used in conjunction with a picture test, might be very accurate in estimating prognosis.

References

1. Lindquist, E. F. *Statistical Analysis in Educational Research.* Boston: Houghton Mifflin, 1940.
2. Roe, V. and R. Milisen. The effect of maturation upon defective articulation in the elementary grades. *JSD*, 7, 1942, 37-50.

VI

THE EFFECTIVENESS OF COMBINED VISUAL-AUDITORY STIMULATION IN IMPROVING ARTICULATION

Davis A. Scott and Robert Milisen

THIS STUDY concerns itself with a more extensive analysis and evaluation of combined visual-auditory stimulation rather than with making comparisons between different forms of stimulation. Likewise, no comparisons will be made with other methods of articulation therapy. The purpose of this study is to attempt to answer the following questions:

1. How effective is a combined visual and auditory stimulation in improving incorrect speech responses of defective speaking children?

2. Are there any measurable differences in the speech responses of children to stimulation at different levels of difficulty in the speech configuration?

3. Are some sounds more stimulatable than others?

Procedure

Subjects. The group of subjects used in this study was the same group as described in the authors' preceding article.

Tests. The picture articulation test and a stimulation test were administered. For the purpose of this study, the stimulation test was administered to determine how well the subject could produce each of his defective sounds in isolation, in nonsense syllables, and in words in response to the examiner's visual-auditory stimulation.

1. For the sound in isolation, the subject was asked to produce the sound following a visual-auditory stimulus; the subject imitated what he heard and saw.

2. For the sound in nonsense syllables, the subject responded to combined visual-auditory stimulation in first the initial position, then in the medial position, then in the final position. In each instance, the vowel [ɑ] was coupled with the sounds.

3. For the sound in words, the subject responded to combined visual-auditory stimulation in first the initial position, then in the medial position, and last in the final position.

The number of stimulations presented depended upon the adequacy of the subject's response. Each sound in each condition described above was stimulated once and the results recorded if the subject's response was correct. If the response was incorrect, a second stimulus was presented and, if necessary, a third one. If none of the imitations were judged as correct, the best of the three responses was recorded.

The examiner tested and stimulated the same sounds, used the same introductory procedures, gave the same directions, and recorded responses in the same manner as described in the preceding article.

Compilation of the Data. In order to make the data as meaningful as possible, two methods of reporting were used: a weighting scale and a right-versus-wrong classification. The method used in assigning weight or score values to the speech responses followed Table 1, the rationale for such a weighting being based on the

TABLE 1. Weight values assigned to recorded responses.

Recorded Responses	*Interpretation*	*Weight Value*
1	correct	1
2 and 3	mild distortions	2
4 and 5	serious distortions	3
[ ]	substitutions	4
0	omissions	5

study by Roe and Milisen (*1*), and described by Snow and Milisen, above.

The technique involved in determining right versus wrong responses was as follows: A response was classified as incorrect if it were an omission, substitution, or distortion. A response was classified as correct if the correct sound had been produced or if the response had been recorded as 2, indicating a mildly indistinct sound which would not be recognized as an error by laymen.

Results

1. *How effective is a combined visual and auditory stimulation in improving incorrect speech responses of defective speaking children?* In answering this question, a comparison was made between the incorrect responses to the articulation test and the responses to the stimulation test. The answer to this question was approached in two ways.

The first way was to determine the degree of significance of the differences which were observed between responses on the articulation test and those on the stimulation test. All of the responses were first translated into weight values according to Table 1. It was then possible to obtain mean score values and to compute numerical differences. In order to determine the significance of these differences, sig-

TABLE 2. Significance ratios for differences between sound responses on articulation test and sound responses in words on stimulation test.

Sound	*N (Subjects)*	*Initial Position*			*Medial Position*			*Final Position*		
		MX / MY*	*Diff.*	*t*	*MX* / MY*	*Diff.*	*t*	*MX* / MY*	*Diff.*	*t*
f	33	3.12 / 1.58	1.54	5.70	3.64 / 1.67	1.97	8.96	4.12 / 1.30	2.82	10.44
v	34	3.97 / 1.77	2.20	9.66	3.79 / 2.18	1.61	5.03	4.12 / 2.35	1.77	5.71
k	37	3.08 / 1.33	1.75	6.20	2.89 / 1.70	1.19	4.25	4.11 / 1.59	2.52	9.33
g	31	3.10 / 1.58	1.52	4.00	3.13 / 1.77	1.36	4.39	3.97 / 2.39	1.58	4.65
r	27	3.22 / 2.33	.89	4.05	3.22 / 2.45	.77	3.85	4.07 / 3.00	1.07	3.24
l	40	3.37 / 1.50	1.87	7.48	4.05 / 2.17	1.88	7.23	4.35 / 2.60	1.75	5.65
s	45	3.47 / 1.96	1.51	6.86	3.62 / 2.49	1.13	5.38	3.80 / 1.93	1.87	7.48
z	35	3.57 / 1.57	2.00	8.70	3.89 / 2.91	.98	3.92	3.94 / 1.57	2.37	8.17

*MX—Mean score on articulation test; MY—Mean score on stimulation test.

TABLE 3. Correct responses following stimulation of sound in initial, medial, and final positions of words when particular sound was defective in corresponding positions of words on articulation test.

Sound	*Initial Position*			*Medial Position*			*Final Position*		
	*N**	*n*†	%‡	*N**	*n*†	%‡	*N**	*n*†	%‡
f	25	19	76.0	29	19	65.7	31	25	80.6
v	35	25	71.4	32	18	56.3	31	17	54.8
k	27	22	81.5	22	12	54.5	32	22	68.8
g	24	15	62.5	23	14	60.9	28	13	46.4
r	26	10	38.5	25	8	32.0	26	6	23.1
l	33	24	72.7	35	18	51.4	38	18	47.4
s	38	25	65.8	37	13	37.8	38	23	60.5
z	34	21	61.8	36	11	30.6	30	21	70.0

*N—Number of subjects producing sound incorrectly in initial (or medial or final) position of words on articulation test.

†n—Number of subjects producing sound correctly in initial (or medial or final) position of words on stimulation test.

‡%—Per cent of subjects producing sound correctly following stimulation after producing sound incorrectly on articulation test.

nificance ratios were computed; the ratio was obtained by dividing the mean of the differences by the standard error of the mean differences. Table 2 summarizes the tabulations and contains the significance ratios. Since all of the ratios are greater than 2.58, the observed differences are all significant beyond the 1% level of confidence. Hence, visual and auditory stimulation is extremely effective in improving incorrect speech responses which appear on articulation tests.

The second method of approach was to determine the frequency of correct responses on the stimulation test for those subjects who had produced sounds incorrectly on the articulation test. The results of this analysis are contained in Table 3. Once again, the effectiveness of visual-auditory stimulation can easily be seen.

2. *Are there any measurable differences in the speech responses of children to stimulation at the different levels of difficulty in the speech configuration?* Comparisons were made between different levels of the stimulation test; restated, the question would be: Are more correct responses found following stimulation of a sound in isolation than following stimulation of that sound in nonsense syllables or words? And, does a sound in nonsense syllables respond better to stimulation than it does in words? Furthermore, do sounds in certain positions respond better to stimulation than they do in other positions?

The method used in answering these questions was to compute the percentage of correct responses at various levels of the stimulation test. The results are reported in Table 4.

With but one exception, [l], there

TABLE 4. Per cent of correct responses on stimulation test at various levels of difficulty in speech configuration.

Sound	*Isolation*	*Nonsense Syllables*			*Words*		
		Initial	*Medial*	*Final*	*Initial*	*Medial*	*Final*
f	100	88	79	84	76	66	81
v	100	94	77	58	71	56	55
k	95	89	77	81	82	55	69
g	91	88	83	71	63	61	46
r	66	39	20	23	39	32	23
l	83	91	94	71	73	51	47
s	91	68	54	55	66	38	61
z	90	71	58	53	62	31	70

were greater percentages of correct responses following stimulation of a sound in isolation than there were at any other level of the stimulation test. Except for one sound, [r], there were greater percentages of correct responses in the initial position of nonsense syllables than there were in the initial position of words. Except for the [r] sound, there were greater percentages of correct responses in the medial position of nonsense syllables than there were in the medial position of words. Except for three sounds, [r s z], there were greater percentages of correct responses in the final position of nonsense syllables than there were in the final position of words. Except for the [f] and [z] sounds, there were greater percentages of correct responses in the initial position of words than there were in the medial or final positions of words. For the [f k s z] sounds, there were greater percentages of correct responses in the final position of words than there were in the medial position; for the sounds [v g r l] there were greater percentages of correct responses in the medial position of words than there were in the final position.

3. *Are some sounds more stimulatable than others?* For the answer to this question, the writer tabulated the total number of possible responses and the total number of correct responses on the stimulation test. This was done for each sound being studied and in-

TABLE 5. Per cent of correct responses out of total number of responses on stimulation test.

Sound	*Total Responses*	*Total Correct Responses*	*Per Cent Correct Responses*
f	205	169	83
v	232	172	74
k	201	160	79
g	184	133	72
r	186	64	34
l	254	185	72
s	268	166	62
z	240	150	63

cluded the responses of each subject. Then the percentage of correct responses out of total responses was computed. The results appear in Table 5.

From this analysis of the data in terms of the per cent of correct responses following stimulation, the ranking of the sounds according to stimulatability would be as follows: [f k v g l z s r]. This order of ranking places at the top sounds with visible focal articulation points along with those sounds with focal articulation points that can be made more visible. Lower in rank appear those sounds with focal articulation points that are relatively non-visible and cannot be made more visible. The one exception to this generalization is the [r] sound with a focal articulation point that can be made more visible; the [r] ranks as the least stimulatable.

Summary

The purpose of this study has been to obtain some information which would point to the effect of visual and auditory stimulation upon the speech responses of defective speaking children. The following summary attempts to answer the questions posed at the beginning of the study.

1. Combined visual and auditory stimulation was extremely effective in improving the incorrect speech responses of defective speaking children. The differences observed between the responses on the articulation test and the responses following stimulation were significant for all the sounds studied and in all positions (Table 2).

2. A high percentage of children with defective articulation were able to produce correct sound responses in all three positions of words following a minimal amount of visual and auditory stimulation. Following visual-auditory stimulation of a sound in the three positions in words, from 23.1% to 81.5% of the responses were correctly articulated sounds (Table 3).

3. Sounds were made correctly more often in isolation than in any position of nonsense syllables or words. There was one exception to this conclusion: The (l) sound was produced correctly more often in the initial and medial positions of nonsense syllables than in isolation. Sounds usually were made correctly more often in any position of nonsense syllables than in the corresponding position of words (Table 4).

4. In nonsense syllables, sounds were produced correctly more often in the initial position than in the medial or final positions and more often in the medial position than in the final position (Table 4).

5. In words, sounds were produced correctly more often in the initial position than in the medial or final positions except for the [f] and [z] sounds. In comparing the medial and final positions of words, some sounds were produced correctly more often in the medial position, [g r l]; other sounds were produced correctly more often in the final position, [z s k f] (Table 4).

6. Sounds with visible focal articulation points and sounds with focal articulation points that can be made more visible were the most stimulatable. An obvious exception to this conclusion was the [r] sound, which appeared to be the least stimulatable, although its focal articulation point can be made more visible (Table 5).

Reference

1. Roe, V. and R. Milisen. The effect of maturation upon defective articulation in elementary grades. *JSD*, 1942, 7, 37-50.

VII

A STUDY OF THE ABILITY TO REPRODUCE UNFAMILIAR SOUNDS WHICH HAVE BEEN PRESENTED ORALLY

William R. Humphrey and Robert Milisen

THE SPEECH therapist is constantly confronted with problems in his work due to a limited knowledge of the effect which memory and mode of presentation have upon the learning of speech sounds. Articulation therapy is an attempt to overcome some of the barriers in the way of accurate speech sound acquisition. There have been many psychological experiments concerning memory and mode of presentation, but they have not dealt directly with the learning of correct articulation.

The purpose of this study was to determine the effect upon the production of unfamiliar sounds of (a) visual cues, (b) auditory cues, (c) visual combined with auditory cues, and (d) frequency of stimulations received by the subject.

Normal speakers, not defective speakers, were used in this study. The justification of this procedure is based on the assumption that articulation therapy largely involves the *teaching* of unfamiliar sounds. There is little reason to believe that the principles involved in learning are different for the normal speaker than for the defective speaker. It may be argued that some defective speakers will have acquired extraneous 'emotional' variables involving attitudes and learning which may interfere with his learning to speak. Although this statement is true it should be pointed out that this study does not compare the learning speed and skill of the normal speaker with those of the defective speaker, but rather compares the ability of the normal speaker to learn unfamiliar sounds under one condition as compared with his ability to learn under another condition. In this way the importance of each condition used in stimulation therapy can be isolated and compared by excluding the uncontrollable variables.

From this study an attempt will be made to answer the following questions:

1. Is one of the stimulus cues (visual, auditory, visual plus auditory) more effective in eliciting sounds?

2. Is the effectiveness of eliciting a sound related to the inherent visibility of the focal articulation point of each sound?

3. Are sounds which are more easily elicited also more readily retained?

4. Is the re-learning of a sound after a long period of non-stimulation easier than the initial learning of that sound? If so, what is the importance of the nature of the sounds and the stimulus cues?

5. Can sounds be learned from stimulation if no reinforcement is given?

Subjects

The subjects for this study were chosen from freshman speech and education classes at Indiana University. Forty-five were female and 15 were male. It was impossible to include Negro subjects, since the lights used in the experimental situation did not provide enough illumination to allow the observer to make good judgments of visible movements.

None of the subjects used had had previous therapy in speech or any training in speech pathology. None of the subjects had speech defects or hearing losses. All of the subjects were freshmen or first semester sophomores and ranged in age from 18 to 29; the mean age was 22.2. The 60 subjects used were randomly divided into three groups of 20 each.

Sounds

Twenty-seven sounds which do not occur in the English language were employed in this study. The sounds were labial, lingual, and guttural noises which were either devised by the experimenter or found in foreign languages used infrequently in the United States. Seven subjects were asked to imitate all of the sounds in a pilot study. Eighteen sounds were excluded because the subjects could not imitate them at all or because they could be imitated correctly on the first trial.

The nine sounds retained were grouped into three categories depending upon visibility of their focal articulation points. Three had visible focal articulation points; three had partially visible focal articulation points; and three had non-visible focal articulation points. The following is a description of each sound:

Visible. 1. Unvoiced, bilabial, friction sound produced by compressing the lips and forcing air between them.

2. Voiced, upper labial, lower dental sound produced by exploding the air through an opening between the teeth and lip.

3. Voiced, bilabial, friction sound produced by blowing a stream of air between lips which were approximated.

Partially Visible. 4. Medial lingual, hard palate, explosive click produced by sucking air into the oral cavity as the tongue is quickly lowered.

5. Voiced, lingual, alveolar, friction sound produced by folding the tip of the tongue under and forcing the top of the tongue against the alveolar ridge and teeth. The air forced between this focal articulation point sounded like a distortion between the [z] and [ʒ].

6. Voiced, lingual, medial hard palate, friction sound produced by flattening the tongue parallel with the hard palate and about ¼-inch behind the alveolar ridge. The air stream was forced between a loose focal articulation point.

Non-visible. 7. Voiced, posterior lingual, soft palate, explosive sound produced by forcing the back portion of the tongue against the soft palate and quickly lowering it.

8. Unvoiced, soft palate, friction sound produced by forcing air against a relaxed soft palate. The sound, produced by the vibration of the velum, passes through the nasal and the oral cavities.

9. Posterior lingual, pharyngeal, explosive sound produced by sucking air inward by enlarging the pharynx while the tongue breaks contact with the pharynx.

The numbers that are used here to designate the sounds are also used throughout the study.

Room and Lighting

The room used was 5 feet by 7 feet. On a small table in the center of the room was mounted a window 16 inches by 21 inches fitted on both sides with regular 14 gauge galvanized screen wire. Surrounding this window and dividing the room was a heavy dark curtain. During the experiment, the experimenter sat on one side of the window and the subject on the other. On the subject's side was a 150 watt light which burned constantly during the experiment. On the experimenter's side was a fluorescent lamp (15 watt cool white) with an accessible switch. This lamp was mounted across the top of the window and was turned on and off depending upon the visibility desired. When the light was out, vision through the screen from the subject's side was negligible. When the light was on, the subject had adequate visibility through the screen. At the bottom of the window on the experimenter's side was a mirror, 9 inches by 21 inches. This mirror reflected the experimenter's image, making it easier for him to control the facial movements used in producing the nine sounds.

Definition of Terms

This study involves terms for stimulus cues and response conditions which are so similar that careful definition is necessary. The stimulus cues involve the experimenter's saying sounds and the response conditions involve the subject's imitation of the experimenter.

The stimulus cues must be divided into two parts. First, the amount of visibility inherent in the focal articulation point used in producing the sound. Thus the sounds are labeled as visible, partially visible or non-visible according to the visibilities of each focal articulation point. Second, the kind of stimulus (visual or auditory) given to the subject by the experimenter. In the stimulation situation the subject received *visual cues* when he could see the movements used in making a sound but could not hear any sound. In the situation where the subject received only *auditory cues*, he could hear the sound but could not see the movements used in making it. When *visual plus auditory cues* were given, the subject was able to see the movement and hear the sound as made by the experimenter.

The responses of the subject were evaluated on the basis of, first, their visible characteristics (visual responses) and, second, their audible characteristics (audible responses).

Procedure

At the beginning when the subject was seated, both lights were on and there was clear vision through the screen. The subject was given these instructions:

> This is a study of your ability to imitate the sounds which I make. Some people can imitate them and some people cannot. Some sounds you will be able to hear, and some sounds you will only see. Some sounds you will see and hear. You will have 10 trials for each sound. In each trial I will present the sound three times and you are to give one response. You will not be told how well you do, and it will be up to you to decide which responses are the best and then continue making these or modifying them as you see fit. Modify your responses after each trial until you think you have achieved the correct response. This does not tax your intelligence nor does it test it. Remember, after hearing the sound three times you respond only once. Are there any questions?

In order that any confusion concerning the procedure might be cleared before the stimulation began, an example was given for each stimulus cue. The example sound was [p]. Each subject was asked to observe what a [p] sound would be like for the three stimulus cues: (a) visual movements only, (b) auditory sounds only, and (c) visual plus auditory stimulation combined. The subject was given ample time to examine the procedure and ask questions.

As soon as the subject understood the procedure, the observer began the experiment by producing the sound three times and recording his judgment of the subject's response. Only the first response was evaluated if the subject produced two or more responses in rapid succession.

TABLE 1. Comparison of the responses to the stimulus cues.

Visibility of Focal Articulation Point	*Audible Responses According to Type of Stimulation*	*Mean*	*Differences in the Audible Responses According to the Type of Stimulation*	*t-score*
Visible Sounds				
1	Visual plus Auditory	23.10	Visual plus Auditory vs. Auditory	1.935
	Auditory	18.80	Visual plus Auditory vs. Visual	11.015*
	Visual	29.25	Visual vs. Auditory	6.411*
2	Visual plus Auditory	14.60	Visual plus Auditory vs. Auditory	1.086
	Auditory	12.60	Visual plus Auditory vs. Visual	7.614*
	Visual	27.85	Visual vs. Auditory	11.507*
3	Visual plus Auditory	14.85	Visual plus Auditory vs. Auditory	1.619
	Auditory	17.15	Visual plus Auditory vs. Visual	10.128*
	Visual	26.70	Visual vs. Auditory	5.777*
Partially Visible Sounds				
4	Visual plus Auditory	16.85	Visual plus Auditory vs. Auditory	.153
	Auditory	17.15	Visual plus Auditory vs Visual	9.570*
	Visual	29.10	Visual vs. Auditory	6.616*
5	Visual plus Auditory	19.20	Visual plus Auditory vs Auditory	1.220
	Auditory	17.15	Visual plus Auditory vs. Visual	4.383*
	Visual	25.60	Visual vs. Auditory	5.559*
6	Visual plus Auditory	20.45	Visual plus Auditory vs. Auditory	1.413
	Auditory	18.50	Visual plus Auditory vs. Visual	6.384*
	Visual	28.75	Visual vs. Auditory	9.403*
Non-Visible Sounds				
7	Visual plus Auditory	15.05	Visual plus Auditory vs. Auditory	.084
	Auditory	15.25	Visual plus Auditory vs. Visual	7.615*
	Visual	29.90	Visual vs. Auditory	10.772*
8	Visual plus Auditory	20.95	Visual plus Auditory vs. Auditory	1.684
	Auditory	17.70	Visual plus Auditory vs. Visual	5.212*
	Visual	28.55	Visual vs. Auditory	7.363*
9	Visual plus Auditory	21.80	Visual plus Auditory vs. Auditory	.201
	Auditory	21.45	Visual plus Auditory vs. Visual	3.928*
	Visual	28.40	Visual vs. Auditory	6.556*

	Visible Responses According to Type of Stimulation	*Mean*	*Differences in the Audible Responses According to the Type of Stimulation*	*t-score*
Visible Sounds				
1	Visual plus Auditory	14.60	Visual plus Auditory vs. Auditory	1.795
	Auditory	18.15	Visual plus Auditory vs. Visual	2.009*
	Visual	16.75	Visual vs. Auditory	.700
2	Visual plus Auditory	18.25	Visual plus Auditory vs. Auditory	5.256*
	Auditory	28.50	Visual plus Auditory vs. Visual	.227
	Visual	17.50	Visual vs. Auditory	6.111*
3	Visual plus Auditory	13.85	Visual plus Auditory vs. Auditory	7.484*
	Auditory	26.05	Visual plus Auditory vs. Visual	.539
	Visual	14.80	Visual vs. Auditory	6.147*
Partially Visible Sounds				
4	Visual plus Auditory	15.50	Visual plus Auditory vs. Auditory	2.488*
	Auditory	20.70	Visual plus Auditory vs. Visual	2.402*
	Visual	19.80	Visual vs. Auditory	.443
5	Visual plus Auditory	16.95	Visual plus Auditory vs. Auditory	7.680*
	Auditory	29.70	Visual plus Auditory vs. Visual	1.243
	Visual	14.60	Visual vs. Auditory	5.729*
6	Visual plus Auditory	19.10	Visual plus Auditory vs. Auditory	4.251*
	Auditory	26.20	Visual plus Auditory vs. Visual	.693
	Visual	17.70	Visual vs. Auditory	4.473*

*Significant differences at the 5% level of confidence or better.

When the subject was given visual plus auditory stimulation, the lights in both parts of the room were turned on. As soon as the subject began to respond, the experimenter's side was darkened so the subject couldn't see him. This eliminated any chance of the subject being reinforced through the medium of vision.

When the subject was given visual stimulation, the lights in both parts of the room were turned on so the subject could see the experimenter make the movement although he could not hear any sound. As soon as the subject began to respond, the experimenter's light was turned out.

When the subject was given auditory stimulation, the light on the experimenter's side was turned out during the time of stimulation and during the subject's responses.

Each subject was stimulated with all nine sounds, three sounds by each stimulus cue: visible, audible and visible-audible. In the three randomly selected groups of subjects, the order of presentation of both the sounds and the stimulus cues was rotated in such a way that each sound was presented an equal number of times by each stimulus cue.

TABLE 2. Relation between the visibility of the focal articulation point and the responses.

Stimulus Cue	*Type of Response*	*Visibility of Focal Articulation Point*	*Mean*	*Comparison*	*t-score*
Visual plus Auditory	Audible	Visible	17.48	Visible vs. Partially Visible	1.216
		Partially Visible	18.83	Visible vs. Non-Visible	1.578
		Non-Visible	19.28	Partially Visible vs. Non-Visible	.368
Visual	Audible	Visible	27.93	Visible vs. Partially Visible	.169
		Partially Visible	27.81	Visible vs. Non-Visible	1.568
		Non-Visible	28.95	Partially Visible vs. Non-Visible	2.033*
Auditory	Audible	Visible	16.18	Visible vs. Partially Visible	1.692
		Partially Visible	17.60	Visible vs. Non-Visible	2.060*
		Non-Visible	18.16	Partially Visible vs. Non-Visible	.528
Visual plus Auditory	Visible	Visible	15.56	Visible vs. Partially Visible	1.384
		Partially Visible	17.18		
Visual	Visible	Visible	16.35	Visible vs. Partially Visible	.821
		Partially Visible	17.36		
Auditory	Visible	Visible	24.23	Visible vs. Partially Visble	1.092
		Partially Visible	25.53		

*Significant at the 5% level of confidence or better.

Each of the subject's responses were evaluated in two ways: (a) visible responses and (b) audible responses. The degree of correctness of each was rated as (1) correct, (2) mild error, (3) marked error. A visible response was rated as a mild error if the correct articulators were used but in not quite the required way. It was a marked error if the wrong articulators were used or if the correct ones were used in a markedly different manner. The evaluation of correctness of audible responses was somewhat similar to the evaluation of errors in articulation testing. The 'mild errors' were similar to slight distortions. The 'marked errors' were similar to severe distortions and/or substitutions, according to Milisen (3).

Situation A was the testing situation. The subject was given 30 stimulations for each sound, responding once after each three stimulations or a total of 10 times for each sound. This procedure was followed for each of the nine sounds.

Situation B was the re-testing situation which occurred 14[1] days after Situation A. The subjects were asked to produce without stimulation each of the sounds they could remember from the previous testing situation. The results were recorded. After this the procedure of Situation A was repeated.

Results

Reliability. Twenty of the subjects were also judged by another observer. The judgments of the experimenter and the observer were correlated for the first 10 subjects. The Pearson product-moment correlation was used. The correlation of the evaluations of the visible responses showed an $r=.787$

[1]The interval of time between testing situations was derived from a study of memory by Ebbinghaus (*1*).

and the evaluations of the audible responses showed an $r = .841$.

Ten additional subjects were studied after the observer had received additional practice. The correlation of the evaluations of visible responses showed an $r = .874$, and the evaluations of the audible responses showed an $r = .909$. Thus, consistency increased with training. In every situation the evaluations of audible responses correlated more highly than visible responses. This may have been due to the room situation. It was not possible for the observer to have as straight a line of vision as did the experimenter.

In analyzing the data and reporting the results, an attempt was made to answer the following five questions:

1. *Is one of the stimulus cues (visual, auditory, or visual plus auditory) more effective in eliciting sounds?* Data were used from Situation A to answer this question (Table

TABLE 3. Production of sounds from memory after an interval of 14 days. Situation B.

		Visible Responses			*Audible Responses*		
Stimulus Cue	*Sound*	*% Correct*	*% Incorrect*	*% No Response*	*% Correct*	*% Incorrect*	*% No Response*
Visual plus Auditory	1	50	5	45	25	30	45
	2	0	20	80	0	20	80
	3	5	10	85	0	15	85
	4	25	35	40	40	20	40
	5	5	15	80	0	20	80
	6	0	10	90	0	10	90
	7				5	20	75
	8				0	10	90
	9				10	40	50
Visual	1	15	20	65	0	35	65
	2	5	5	90	0	10	90
	3	25	10	65	0	35	65
	4	0	25	75	0	25	75
	5	15	10	75	0	25	75
	6	5	15	80	0	20	80
	7				5	10	85
	8				0	20	80
	9				0	10	90
Auditory	1	30	20	50	25	25	50
	2	0	15	85	5	10	85
	3	0	20	80	5	15	80
	4	10	40	50	25	25	50
	5	0	20	80	0	20	80
	6	10	20	70	0	30	70
	7				20	20	60
	8				20	20	60
	9				5	20	75

TABLE 4. Comparison of the order of sounds in situation A and their order of retention.

Stimulus Cue	Audible Response: Sound	Audible Response: Rank In A	Audible Response: Rank in Retention	Visible Response: Sound	Visible Response: Rank In A	Visible Response: Rank in Retention
Visual	1	9	1	1	2	1
plus	2	1	6	2	5	5
Auditory	3	2	7	3	1	3
	4	4	2	4	3	2
	5	5	5	5	4	4
	6	6	8	6	6	6
	7	3	4			
	8	7	9			
	9	8	3			
	Correlation p = –.26			Correlation p = .83		
Visual	1	8	1	1	3	3
	2	3	7	2	4	4
	3	2	2	3	2	1
	4	7	3	4	6	6
	5	1	4	5	1	2
	6	6	5	6	5	5
	7	9	8			
	8	5	6			
	9	4	9			
	Correlation p = –.14			Correlation p = .85		
Auditory	1	8	1	1	1	1
	2	1	4	2	5	5
	3	3	5	3	3	4
	4	4	2	4	2	2
	5	5	9	5	6	6
	6	7	7	6	4	3
	7	2	3			
	8	6	8			
	9	9	6			
	Correlation p = .21			Correlation p = .94		

1). It is apparent from these results that when attempting to elicit an audible response, there was very little difference between visual plus auditory stimulation and auditory stimulation. When eliciting visible responses, there was very little difference between visual plus auditory stimulation and visual stimulation.

Good visible responses were elicited by both auditory stimulation and visual stimulation while audible responses were not elicited very well by visual stimulation used alone.

2. *Is the effectiveness of eliciting a sound related to the inherent visibility of the focal articulation point of the sound?* Data were used from

TABLE 5. Comparison of the results in situations A and B.

Type of Response	*Stimulus Cue*	*Visibility of Sound*	*Mean Situation A*	*Mean Situation B*	*Difference Between A and B*	*t-score*
Audible	Visual	Visible	17.48	15.91	1.57	1.42
	plus	Partially Visible	18.83	15.35	3.48	2.80 *
	Auditory	Non-Visible	19.26	17.58	1.68	1.27
Audible	Visual	Visible	27.93	26.30	1.63	2.01 *
		Partially Visible	27.81	26.63	1.18	1.55
		Non-Visible	28.95	27.70	1.25	2.01 *
Audible	Auditory	Visible	16.18	15.15	1.03	1.18
		Partially Visible	17.60	16.45	1.15	1.07
		Non-Visible	18.16	18.03	.13	.108
Visible	Visible	Visible	15.56	12.73	2.83	2.46 *
	plus	Partially Visible	17.18	14.91	2.27	2.46 *
	Auditory					
Visible	Visual	Visible	16.34	11.56	4.78	5.20 *
		Partially Visible	17.13	16.2	.93	.815
Visible	Auditory	Visible	24.23	22.50	1.73	1.57
		Partially Visible	25.53	23.85	1.68	1.23

*Significant at the 5% level of confidence or better.

Situation A to answer this question (Table 2). Although the differences obtained were not highly significant, a consistent trend indicated that the highly visible sounds were the easiest to elicit. The trend also indicated that the partially visible sounds were elicited more easily than non-visible sounds.

3. *Are sounds which are more easily elicited also more readily retained?* Data were used from the over-all results of Situation A (Table 1) and the first part of Situation B (Table 3). The order of learning of the sounds was compiled and the order of retention of the sounds was compiled. These distributions were compared according to their rank correlation (2). The results appear in Table 4.

These results indicate a definite trend. The visible responses of the sounds with the highest correct responses in Situation A were retained best. The visible responses with the fewest correct responses in Situation A were retained least. No significant correlation was present in the audible responses.

4. *Is the re-learning of a sound after a long period of non-stimulation easier than the initial learning of that sound? If so, what is the importance of the nature of the sounds and the stimulus cues?* Data were used from Situation A and Situation B to answer this question (Table 5). For every sound and under every condition there was an increase in correct responses in Situation B over Situation A, although in many instances the differences were not significant. Visible responses showed a much greater significant increase in learning sounds in Situation B than in Situation A. The

TABLE 6. Percentage of correct audible responses for stimulus trial.

	Stimulus Cues								
	Visual plus Auditory			*Visual*			*Auditory*		
Stimulus Trials	*Visible*	*Partially Visible*	*Non-Visible*	*Visible*	*Partially Visible*	*Non-Visible*	*Visible*	*Partially Visible*	*Non-Visible*
Situation A									
1	33	22	31	1	0	0	38	23	18
2	30	28	30	3	0	0	50	33	21
3	42	28	31	5	0	0	48	25	25
4	42	28	33	5	0	0	53	35	30
5	46	33	36	5	0	0	43	35	30
6	52	28	40	5	0	0	51	38	25
7	50	33	40	6	0	0	51	40	28
8	53	31	36	3	0	0	50	36	28
9	60	33	40	5	0	0	48	36	36
10	57	31	41	3	0	0	46	38	33
Situation B									
1	46	41	40	1	1	3	51	31	31
2	41	45	43	3	1	5	56	33	31
3	53	53	46	5	6	5	55	38	28
4	53	60	46	5	5	3	55	48	43
5	54	54	50	3	6	5	58	50	36
6	61	60	45	5	3	5	61	45	36
7	63	61	45	3	3	3	65	45	36
8	66	60	45	3	3	3	61	45	33
9	63	60	46	3	3	1	60	48	38
10	66	53	45	3	3	1	53	48	38

visible response to visible sounds stimulated by visual cues made the greatest increase between the two situations. These results tend to be comparable to those found in answering question number 3.

In comparing the sounds it was found that the visible sounds made the greatest increase in responses between the two situations. Partially visible sounds gave greater differences than non-visible sounds.

In comparing the stimulus cues for all responses, visual plus auditory cues tend to give greater differences between the two situations than visual or auditory cues. Visual cues tend to give greater differences than auditory cues.

5. *Can sounds be learned from stimulation if no reinforcement is given by the experimenter?* Data were used from Situations A and B to answer this question (Tables 6 and 7). The percentage of correct audible and visible responses was determined for all subjects under all stimulus conditions in Situations A and B. The percentages indicated an increase in correct responses for every sound. Learning took place since the frequency of correct responses increased somewhat in proportion to the amount of stimulation given, even though the responses were not reinforced by the experimenter. It should be noted, however, that the responses may not have been as stable as they might

TABLE 7. Percentage of correct visible responses per stimulus trial.

	Stimulus Cues					
	Visual plus Auditory		Visual		Auditory	
Stimulus Trials	Visible	Partially Visible	Visible	Partially Visible	Visible	Partially Visible
Situation A						
1	46	26	38	35	13	6
2	53	38	43	30	13	8
3	56	41	53	35	13	8
4	65	45	50	36	15	13
5	61	48	53	40	15	15
6	60	48	53	43	15	13
7	61	50	56	43	15	15
8	61	51	53	41	15	16
9	61	48	51	41	15	16
10	65	43	53	31	15	15
Situation B						
1	83	46	88	50	33	10
2	85	46	86	53	26	11
3	85	48	91	53	25	8
4	78	48	88	56	28	15
5	80	48	86	51	28	13
6	78	51	86	50	28	11
7	78	53	90	56	26	13
8	76	53	88	53	26	16
9	75	60	90	51	28	18
10	78	53	90	51	23	20

have been with reinforcement in light of the fact that the percentage of correct responses increased to a peak and then fell off slightly. The adequacy of the last responses regressed somewhat but were better than the first responses had been.

Conclusions

The design of this study appears to offer a practical method of studying the principles involved in the learning of unfamiliar sounds. In so far as articulation therapy involves the learning of unfamiliar sounds by defective speakers, the principles demonstrated by this type of experiment should be directly applicable even though the subjects were normal speakers. It was interesting to note the similarity of response to stimulation of unfamiliar sounds by these normal speakers to the responses made by defective speakers to stimulation of their misarticulated sounds as reported in the studies by Scott (see study by Scott and Milisen, above).

The following principles can be drawn from the relative effectiveness of the kinds of stimulus cues in the reproduction of unfamiliar sounds.

1. *Stimulus cues produced by the experimenter:* Visual plus auditory stimulation resulted in significantly better visible and audible responses

more consistently than either visual or auditory cues alone. Visual cues produced significantly better visible responses than did auditory cues. Auditory cues produced significantly better audible responses than did visual cues. Visual cues tended to produce better audible responses than auditory cues produced visible responses. In terms of improvement in Situation B, results were more significant for visual plus auditory cues than for either visual or auditory cues alone.

2. *Stimulus cues inherent in the sound:* The results indicate a trend that sounds with highly visible focal articulation points were most effective in producing correct responses. Sounds with partially visible focal articulation points were next most effective in producing correct responses. Sounds with non-visible focal articulation points were least effective.

In this experiment, however, neither of the kinds of stimulus cues produced significantly better retention. The retention of learned visible responses was proportionate to the accuracy manifested during the learning situation. This was not true for the audible responses. The visible responses tended to be retained better than the audible responses.

Generally, and regardless of the kind of stimulus cues, a long interval of non-stimulation did not adversely affect the learning of unfamiliar sounds. This fact was substantiated in most cases.

Reproduction of unfamiliar sounds improved significantly even though only a simple stimulus was given and the response was not reinforced by the experimenter.

References

1. Ebbinghaus, H. *Memory, A Contribution to Experimental Psychology.* New York: Columbia Univ. Press, 1936.
2. Lindquist, E. F. *Statistical Analysis in Educational Research.* New York: Houghton Mifflin, 1942.
3. Milisen, R. L. Interpretation of articulation test results. *Speech Hear. Therapist,* Feb., 1945, 6-10.

VIII

EFFECT OF LATENCY BETWEEN STIMULATION AND RESPONSE ON REPRODUCTION OF SOUNDS

Edward F. Romans and Robert Milisen

MOST TEXTBOOKS in the field of speech correction mention stimulation as an important factor in speech development (*11*) as well as in correction (*4*, 7, *10*) of defective articulation. In spite of the fact that stimulation prior to the attempt to produce the sound is widely used as a therapeutic procedure, the effect of latency between the stimulus and response in this situation has not been studied experimentally.

Ebbinghaus (*1*) measured retention of nonsense syllables as affected by the lapse of time between stimulation and response. The numerous studies in rote-learning which followed his research showed that a relationship between the time of presentation and recitation does exist. There is general agreement that latency does affect reminiscence, which is defined by McGeoch (*6*) as ' . . . the improvement in recall of incompletely learned material after an interval of time without intervenient formal relearning or review.'

Many studies of reminiscence using a memory drum have been reported. Hull's memory drum (*3*), a typical model in most experiments of this nature, consisted of a revolving drum which exposed each nonsense syllable for a period of two seconds. Schlosberg (*9*) found that a latency of two seconds produced better learning than a latency of one second. Hovland (2) found that fewer trials were needed for learning with a four-second than a two-second latency. He also found that more syllables were remembered after two seconds than upon immediate recall following stimulation. Ward (*12*) using six-second, 30-second, two-minute, five-minute, 10-minute and 20-minute latencies, found that the two-minute latency group needed fewest trials for learning.

Although speech is a learned process experimental work has not isolated the conditions which precipitate misarticulation. This experiment was designed to give information involving the learning of sounds not in American speech. No experimental evidence has been available on the relationship of the time interval between stimulus and response and the effectiveness of articulation therapy. Should the therapist require the child to respond immediately to the stimulation given, or will a period of reminiscence after the stimulation improve the quality of the response?

The purpose of this experiment was to study the effect of varying the length of latent periods on the quality (correctness) of responses produced while learning unfamiliar sounds through imitation.

The task in this study required groups of normal speaking subjects to reproduce sounds not used in American speech and which were thought to be largely unfamiliar. A number of sounds were used in a pilot study and all but six were eliminated either because they were too easy or too hard to reproduce.

The six sounds retained for use in this experiment were all unvoiced. The focal articulation points varied in visibility in that two sounds were readily visible, two were partially visible and two were non-visible.

Sounds and Materials

Subjects having normal articulation were used since speech defectives might have presented numerous uncontrollable variables. The subjects were given a new task to perform, namely learning unfamiliar sounds. These sounds were as follows:

The *labial-lingual-plosive:* produced by projecting the tongue between the teeth and lips. The correct visible position was made when the tongue made contact with the upper and lower lips. The visible position was considered incorrect if the teeth were at all visible. As the tongue was drawn back, air was expelled sharply. The audible response was an unvoiced sound which approximated the [t] sound. This focal articulation point was readily visible.

The *lingual-trill:* the correct visible response was made when the tongue was vibrating against the alveolar ridge while the mouth was open. The correct audible response was an unvoiced trill. The focal articulation point of this sound was readily visible.

The *uvular-fricative:* instead of expiration of air without obstruction as in the [h], the air stream was directed against the uvula, producing a throat-clearing sound. This was the correct audible sound. The focal articulation point was not readily visible since it was too far back in the mouth.

The *forced-labial-fricative:* the lips were pressed together and protruded for the correct visible position. The correct audible response approximated the 'Bronx cheer' or the motor-boat 'putting.' The sound was made by emitting air through the pressed and protruded lips. This focal articulation point was readily visible.

The *lingual-alveolar-click:* the correct visible position was made when the mouth was open and the tongue was pressed against the entire surface of the hard palate. The correct audible response was made by drawing in air while quickly lowering the tongue from the palate. The correct sound was a short, sharp click. This focal articulation point was partially visible.

The *lateral* [s]: instead of the normal frontal emission, air was forced over the lateral aspects of the tongue since the tongue tip was pressed against the alveolar ridge and the medial incisors. The audible response was a lateral lisp. The focal articulation point was non-visible.

The above sounds required some trials before they were learned, but, on the other hand, they were eventually learned by most of the subjects used in the pilot study. Moreover, each of the six sounds chosen for this study had different auditory and visual characteristics, were unvoiced, were not part of a blend, and were not used in American speech.

Each sound was presented to each subject, who, after a time interval reproduced it. The timing and signal device used in regulating the latency between the stimulation of the sound and its reproduction, was a one rpm Telechron synchronous electric motor and three discs any one of which would fit on the drive shaft. Elevations on each disc depressed a switch which opened the signal light circuit for the desired period of time: 3 seconds, 9 seconds, or 27 seconds. At the moment the circuit was opened the light went off, and the sound was presented orally. The subject reproduced the sound at the moment the light circuit was closed and the light came on.

The timing-signal device was used in a different manner when the subjects were required to reproduce the sound immediately after stimulation. The experimenter depressed the switch with his finger just before he presented the sound and released the switch a moment thereafter to signal the subject to make the response.

Procedure

One hundred and twenty junior high school pupils of the Louisville, Kentucky, Western Junior High School were used. Sixty were boys and 60 girls, 40 were seventh graders, 40 eighth graders and 40 ninth graders. Since all testing was done in the morning before classes started children who were available on the playground were chosen. Most of them were asked to participate in the experiment by the experimenter, but a few were asked to come to the testing room by a child who was just previously tested. All subjects had previously received an audiometric test and were found to have normal hearing.

The procedure involved the presentation of six sounds, 10 times each to 120 junior high school pupils divided into four groups. Four groups were used in order to study one variable in the learning situation, the latency between stimulation and response. The groups were the *0-second latency group* composed of 30 children, half male and half female, equally distributed between the seventh, eighth and ninth grades, the *3-second latency group*, the *9-second latency group* and the *27-second latency group*, all composed of the same number, distributed as in the *0-second group*. For the *0-second group* the experimenter presented the sound once, and the child imitated it immediately. The same stimulation procedure was followed for the subjects in the *3-*, *9-* and *27-second latency groups*. The chief difference was the time interval before the subject responded.

When the subject entered the testing room, he was told:

> You are part of an experiment to see whether junior high school children can make six sounds which are not in the English language. When I make the sound for you to imitate, I want you to look at my face and to listen to the way I make the sound. After I make it, I want you to try to do it exactly as I did. It doesn't matter whether or not you can make the sound, this is not a speech test, but each time I make it, you watch and listen and they try to imitate it exactly. Let's practice the [b] and [l] sounds before we start the experiment.

During the practice period the subject was familiarized with the latency aspect. The timing device was turned on. The subject was told to watch and listen to the experimenter as the sound was made so that he could imitate it when the light came on. The experimenter made the [b] sound when the signal light circuit was opened and the subject was told to imitate the sound when the light came on. The same was done with the [l] sound. Both sounds were practiced until the subject understood the procedure and responded correctly to the timing-signal device. For the *3-second latency group* the same procedure was followed; but using the disc with seven elevated sections to control the light, and for the *9-second latency group* and the *27-second latency group* the appropriate discs were used.

The subjects were only trained to wait the required latency period. At no time were the subjects told whether the response was correct or incorrect. In this manner motivation was not intentionally reinforced. The subjects were told that they were taking part in an experiment, but it would not be a part of their school record. Saltzman (*8*) found that when subjects were given material to learn, they were able to do so faster and make fewer mistakes if there was reinforcement by telling them immediately that their answers were correct. Kausler (*5*) found that subjects (ego-involved) were reinforced and did superior work if they were told the test would be part of a permanent record as compared to the group (task-involved) who were told they were part of a learning situation and test marks were unimportant.

The procedure for all sounds was as follows: signal light off; sound presenter; light on; response by the subject; evaluation of visible and audible responses were recorded by experimenter. This same procedure was followed five times for each of the six sounds. The order of presentation of the sounds was (1) *labial-lingual plosive*, (2) *alveolar-click*, (3) *uvular-fricative*, (4) *forced-labial fricative*, (5) *lingual-trill*, and (6) *laterals*. After each sound had been presented five times, the whole procedure was repeated. At the end of the testing period each subject had made a total of 60 responses and had been scored 60 times for visible and 60 times for audible responses.

The procedures for the four groups were the same except for the latency periods of zero seconds, three seconds, nine seconds and 27 seconds. Throughout the experiment a subject was tested for immediate response, the next subject with a three-second latency, the next with a nine-second latency, the next with a 27-second latency and so on until the 120 subjects were tested.

Judgments of the audible responses were made on the basis of the correctness or incorrectness of the sound produced. Judgments of the visible responses were similarly made on the basis of the correctness or incorrectness of the structures at the focal articulation point. For the sounds which had non-visible focal articulation points, the visible response of the face, lips, etc. had to be judged.

Results

A reliability measure was obtained by comparing the experimenter's judgments with those made by a speech therapist from the Louisville Public Schools who received her Master's degree from the University of Wisconsin. All terms were defined, the sounds were practiced, two subjects were scored for practice, and the judgments were discussed. Reliability was then obtained by inde-

TABLE 1. Means and standard deviations of visible and audible responses for males, females and combined scores.

Latency	*Males*		*Females*		*Combined*	
	Mean	*S.D.*	*Mean*	*S.D.*	*Mean*	*S.D.*
Visible Responses						
0	41.46	7.66	38.06	9.42	39.80	7.82
3	36.40	11.23	40.06	8.60	38.20	7.91
9	45.20	3.83	41.40	11.06	43.30	3.52
27	33.60	7.20	30.40	8.23	31.60	9.21
Audible Responses						
0	36.20	10.21	31.20	10.20	33.70	10.27
3	33.40	12.04	30.67	8.60	32.70	8.12
9	42.33	11.10	33.40	11.70	38.00	8.00
27	24.27	12.00	22.53	6.50	23.40	12.40

TABLE 2. Mean differences between groups and *t*-scores of the combined scores.

Latency in seconds	*Visible* Mean diff.	*t-scores*	*Audible* Mean diff.	*t-scores*
0 versus 3	1.60	.065	1.00	.069
0 versus 9	3.50	1.458	4.30	1.791‡
0 versus 27	8.20	3.350*	10.30	4.291*
3 versus 9	5.10	2.112†	5.30	2.487†
3 versus 27	6.60	2.737*	9.30	3.679*
9 versus 27	11.40	4.850*	14.60	6.083*

*Significant at the 1% level of confidence.
†Significant at the 5% level of confidence.
‡Significant at the 10% level of confidence.

pendent scoring of 12 subjects, six boys and six girls. The procedure was the same as in the experiment. Three subjects were used from each of the four groups. A total of 720 responses were scored for visible and for audible responses, and a Spearman-Brown correlation coefficient of .93 ± .06 was obtained for the visible and a .87 ± .09 for the audible responses. These correlations seem to indicate that the procedures followed by the experimenter could also be followed by other trained persons and with considerable consistency.

The results of this experiment were divided into three parts; scores for males, scores for females, and scores for males plus females. When the term 'combined' is used, it refers to correct responses of both male and female of a particular latency group.

Table 1 shows the means and S.D. of visible and audible responses for male, female, and combined scores. Comparing mean responses, the *9-second latency group* reproduced the sounds more accurately than the *0-, 3-, and 27-second latency* groups. The *9-second latency group* mean scores

TABLE 3. Visible response differences between the first 30 responses and the second 30 responses of all groups.

Latency	*Training period*	*Mean*	*S.D.*	*t-scores*
0	1st half	18.0	4.78	5.362*
	2nd half	21.7	3.48	
3	1st half	17.9	4.54	3.455*
	2nd half	20.6	5.21	
9	1st half	19.9	3.41	5.071*
	2nd half	23.5	3.21	
27	1st half	12.8	5.14	5.202*
	2nd half	18.5	4.87	

*Significant at the 1% level of confidence.

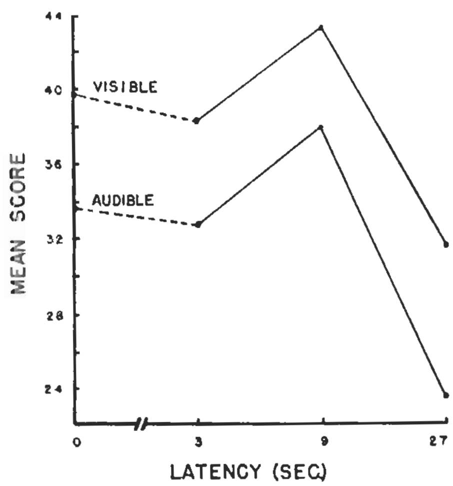

FIGURE 1. The learning curves for all latencies and for visible and audible responses of the combined scores.

were higher for male, female and combined scores in both visible and audible responses.

The males reproduced the six sounds better than the females. Their mean scores were larger for all responses except for visible responses of the three-second latency group. During testing it was evident, however, that the females showed considerable uneasiness. Numerous girls giggled, squirmed in their seats, and manifested a generalized uncomfortable behavior especially in relation to the *forced-labial fricative* or 'Bronx cheer,' and the *uvular-fricative* or throat-clearing sounds. They were disturbed when the experimenter made these sounds as well as when they imitated the sounds. On the other hand, the males seemed to show no such reaction to the two sounds. The results from these two sounds were studied separately, and the boy's scores were compared with the girl's scores. The same was done for the other four sounds. The boys performed better 15.5 per cent of the time on the two sounds which seemed to upset the girls and only two per cent of the time on the other four sounds. Hence the significant difference which was found when all the girls' scores were compared with all the boys' scores, was probably due to the choice of sounds and not to a real difference in learning capacity.

Table 2 shows mean differences between groups and *t*-scores of the combined scores:

1. The *27-second latency group* performed significantly poorer than did the *0-*, *3-*, and *9-second latency groups* for both visible and audible responses.
2. The mean responses following

TABLE 4. Audible response differences between the first 30 responses and the second 30 responses of all groups.

Latency	*Training period*	*Mean*	*S.D.*	*t-scores*
0	1st half	15.0	5.27	4.297*
	2nd half	19.3	5.32	
3	1st half	14.7	4.52	3.890*
	2nd half	18.0	4.75	
9	1st half	17.0	3.96	4.814*
	2nd half	21.0	4.78	
27	1st half	10.0	5.41	5.323*
	2nd half	13.6	6.69	

*Significant at the 1% level of confidence.

the nine-second latency were consistently better than for either the zero, or three-second latency. No significant differences were found, however, comparing zero- with nine-second latency for either visible or audible responses.

3. A significant difference to the five per cent level of confidence occurred for both the visible and audible responses when the *3-second latency group* was compared with the *9-second latency group*. Figure 1 shows the learning curve for all latencies and for visible and audible responses of the combined scores. The learning curve tends to be slightly bimodal for both visible and audible responses with the *9-second* and *0-second latency groups* highest. It was interesting to note that the visible and audible responses described the same pattern.

Tables 3 and 4 compare the results of the first 30 responses with the last 30 responses of all groups. All groups performed significantly better during the last half of their training period than during the first half for visible as well as audible responses.

Conclusions

1. Visual and auditory stimulation can be used to elicit unfamiliar sounds from junior high school children.

2. Learning took place even though no directed reinforcement was given to the reproductions.

3. A differential in response was obtained by changing the latency between stimulation and reproduction.

4. The length of the latency precipitated essentially the same differential for visible as for audible responses, although the audible responses were judged as consistently poorer.

5. The nine-second latency precipitated the best responses. A significant difference was not found, however, between the results of the zero- and nine-second latencies. This condition is important enough to warrant additional investigation.

6. Therapists may profit by allowing more time for reminiscence after stimulation of a sound.

References

1. Ebbinghaus, H. *Memory*. (Trans. by H. A. Ruger and C. E. Bussenius, New York: Teachers Coll., 1913) 1885.
2. Hovland, C. I. Experimental studies in rote-learning theory: II, Reminiscence with varying speeds of syllable presentation. *J. exp. Psychol.*, 22, 1938, 338-353.
3. Hull, M. E. Anticipatory speech responses in children with articulatory defects. *JSHD*, 13, 1948, 268-272.
4. Johnson, W., S. F. Brown, J. F. Curtis, C. W. Edney and J. Keaster. *Speech Handicapped School Children.* New York: Harper, 1948.
5. Kausler, D. H. A study of the relationship between ego-involvement and learning. *J. Psychol.*, 32, 1951, 225.
6. McGeoch, J. A. The influence of four different interpolated activities upon retention. *J. exp. Psychol.*, 14, 1931, 400-413.
7. Melton, A. W. and G. R. Stone. The retention of serial lists of adjectives over short time-intervals with varying rates of presentation. *J. exp. Psychol.*, 30, 1942, 295-310.
8. Saltzman, I. J. Delay of reward and human verbal learning. *J. exp. Psychol.*, 41, 1951, 437-439.
9. Schlosberg, H. An inexpensive memory drum. *J. exp. Psychol.*, 29, 1941, 161-163.
10. Travis, L. E. *Speech Pathology*. New York: D. Appleton, 1931.
11. Van Riper, C. *Speech Correction, Principles and Methods*. (Rev. Ed.) New York: Prentice-Hall, 1947.
12. Ward, L. B. Reminiscence and rote-learning. *Psychol. Monogr.*, 49 (4), 1937.

IX

THE INFLUENCE OF INCREASED STIMULATION UPON THE PRODUCTION OF UNFAMILIAR SOUNDS AS A FUNCTION OF TIME

Donald B. Rice and Robert Milisen

A SURVEY of the literature in speech correction indicates that there are three major techniques used in articulation therapy. Van Riper (*5*) and Curtis (*1*) list these techniques or types of therapy as follows: (a) the phonetic placement method, (b) the moto-kinesthetic method, and (c) the stimulation or stimulus method.

Of these three techniques, the stimulation method is frequently considered to be the most basic and the method most frequently employed in a plan of therapy. According to Curtis (*1*), this method has three distinct advantages: (1) It is the simplest and easiest to use, (2) it is the most direct of all methods, and (3) no distracting stimuli or irrelevant cues are introduced.

The stimulation method is outlined by Travis (*4*), who points out that an important factor in learning is repeated stimulation. This repeated stimulation is often more important than drill, especially when the student is learning to produce a new sound. This author also stated that no experimental evidence was available, but experience in articulation therapy indicated that repeated stimulation over a period of time was frequently necessary to produce the desired response.

The stimulation method begins training on the sound in isolation, then in nonsense syllables, words, phrases, sentences, and finally in reading and propositional speech. While this method has been accepted and used by many speech correctionists, there has been a lack of experimental evidence to aid in determining how often a sound must be stimulated in order to be produced adequately after a lapse of time. Experimental evidence is needed which will give some indication of the effect of the use of additional stimulation with reinforcement upon the production of a sound.

The purpose of this study was to investigate the ability of a person to reproduce unfamiliar sounds as a function of: (1) the amount of stimulation and reinforcement given, and (2) time interval.

Subjects and Material

The subjects used in this experiment were from two sections of 9A English classes at Bloomington High School, Bloomington, Indiana. These classes were chosen because they were available and because English is required for all students, and would therefore tend to provide a random sample of that ninth grade population. Forty-five subjects were used in this study. These subjects made up three groups of 15 each as follows: Group I. 8 females and 7 males; age range 14-3 years to 15-9 years, mean age 14-9 years. Group II. 7 females and 8 males; age range 14-2 years to 16-4 years, mean age 14-8 years. Group III. 8 females and 7 males; age range 14-1 years to 15-9 years, mean age 14-8 years. The members of each group were chosen by the chance order in which they came for training.

Three unfamiliar[1] sounds were used for this study. The sounds were taken from the study by Humphrey, above, who used nine sounds not found in the English language. Of these nine sounds, three were classified as 'visible' because of a visible focal articulation point, and by the same condition, three were classified as 'partially visible' and three as 'non-visible.'

[1]The unfamiliarity of a sound refers to the fact that they are not common speech sounds and would not have been used previously, except by accident or in meaningless babble.

Since the sounds in the English language vary in visibility, it was determined that the three sounds used in this study should show a similar variability of focal articulation point. Therefore, of the three sounds used, one had a highly visible focal articulation point, one a partially visible focal articulation point and the third a non-visible focal articulation point. The three sounds used were described by Humphrey as follows:

> *Visible sound:* Voiced, bilabial, friction sound produced by blowing a stream of air between lips which were approximated.
>
> *Partially visible sound:* Voiced, lingual, alveolar, friction sound produced by folding the top of the tongue against the alveolar ridge and teeth. The air forced between this focal articulation point sounded like a distortion between the [z] and [ʒ].
>
> *Non-visible sound:* Unvoiced, soft palate, friction sound produced by forcing air against a relaxed soft palate. The sound produced by the vibration of the velum, passes through the nasal and oral cavities.

In order to cancel out the possible effect of the order of presentation of the sounds for training, each of the three groups were randomly divided into equal thirds. The first third of each group received training on the three sounds in the following order: visible, partially visible, non-visible. The second third of each group: partially visible, non-visible, visible. The final third of each group: non-visible, visible, partially visible.

Definition of Terms

A *stimulation cycle* is defined as: three productions of the sound by the experimenter, followed immediately by an imitation of the sound by the subject. This imitation was then evaluated by the experimenter and the subject was told whether the response was correct or incorrect. This verbal evaluation of the response constituted the reinforcement given to the subjects in this experiment.

A *stimulation cycle*, therefore, contains three segments: (1) *stimulation*, (2) *imitation*, and (3) *reinforcement*. In this study, the term *stimulation* referred to the first segment of the stimulation cycle. The term *imitation* referred to the second segment and the term *reinforcement* referred to the third segment of the stimulation cycle.

A *production trial* is defined as a measure of retention and consisted of the subject giving 15 consecutive productions of the unfamiliar sound which had previously been stimulated in five or more stimulation cycles. No direct stimulation occurred, however, immediately preceding any of the productions in the production trial.

Procedure

As stated earlier, three groups of subjects were used in this study. Group I represented the *control group*. Group II was made up of those subjects with *slight additional training* and Group III consisted of those with *greater additional training*.

All subjects reported for three experimental sessions. The first session was the only one in which the experimental procedure differed for the three groups. In this session, each of three unfamiliar sounds was presented to each subject. The subjects in Group I received five stimulation cycles for the first sound. This was immediately followed by a production trial of that sound, in which the experimenter recorded on a scoring sheet the value judgments of the 15 productions, but gave no reinforcement. This was called the *immediate* *test situation.* The same procedure was followed for the second sound and then for the third sound.

The procedure for Group II was similar to that of Group I, up to and including the immediate test situation for each sound. At that point, however, the subjects in Group II received the slight additional training on the sound just tested, which consisted of *five* additional stimulation cycles of the sound immediately after the immediate test situation was finished. The same procedure was followed for the other two sounds.

The procedure for Group III was similar to that of Group II, with the following exception: The subjects in Group III received a greater amount of additional training, that is, 15 additional stimulation cycles.

The second experimental session consisted of retesting each subject one hour after the immediate test situation. At this time, each subject gave a production trial of each sound, in the order of presentation in the initial training period. This was called the *one hour test situation.*

The third experimental session consisted of retesting each subject 72 hours after the immediate test situation. At this time, each subject again gave a production trial of each sound, in the order of presentation in the initial training period. This was called the 72 *hour test situation.*

Scoring of Responses

Each response of each subject, during the training period as well as at the time the production trial was given, was scored by the examiner on a four-point scale from three to zero as follows: A judgment of *three* points represented a correct response. A judgment of *two* points was given for a close approximation of the desired sound. A judgment of *one* point was given for a recognized attempt to produce the sound, however severely distorted. A judgment of *zero* was used for a distortion so severe as to be unrecognizable as the desired sound. It was also used to signify a substitution of another sound, or a lack of response because the subject had not retained the sound.

In forming the judgments, the experimenter did not watch the subject, that is, the experimenter looked down at the scoring sheet. This was done so that value judgments would be made on only the audible aspects of the sound, giving no consideration as to whether or not the sound 'looked' correct.

Before the study could be conducted, it was necessary to determine the reliability of the experimenter in making evaluations according to the procedure described and using the sounds which were chosen. For this reliability check, a group of 20 subjects was selected from Study Hall classes at Bloomington High School, Bloomington, Indiana. Another examiner, one who had participated in Wright's study reported above, which concerned the reliability of trained clinicians making evaluations during basic articulation and stimulation testing, was called in to score the responses.

Each of the 20 subjects in this group was given five stimulation cycles on all three sounds. Following the five stimulation cycles for each sound, the subjects gave one production trial for that sound. This procedure was followed for each of the other two sounds. This procedure was the same as that described earlier for the initial training period and the immediate test situation. The examiner and the experimenter scored independently each of the 45 responses of each sub-

TABLE 1. Mean response scores for groups I, II and III, for the immediate, the one hour and the 72 hour test situations.

Group	*Immediate*	*One Hour*	*72 Hours*
I	2.27	.61	.27
II	2.11	.87	.63
III	2.45	1.46	1.08

ject, that is, 15 responses for each of the three sounds. A coefficient of correlation based on the individual scores was used as a determiner of the reliability of the two sets of scores in this situation.

Using a Pearson product-moment coefficient of correlation formula from Lindquist (4), the following coefficients of correlation resulted:

For the visible sound: +.93
For the partially visible sound: +.95
For the non-visible sound: +.83

The results of the reliability measure show considerable consistency between the results of the two examiners. It was concluded that the sounds chosen, the method of evaluat-

TABLE 2. The means of and the difference between the means of groups I, II and III, for the immediate, one hour and 72 hour test situations. (N=15)

	Mean	*S.D.*	*Mean diff.*	*S.E. diff.*	*t*	*Sig. level*
Immediate						
Group I	2.27	.565	.16	.22	.727	n.s.
Group II	2.11	.586				
Group I	2.27	.565	–.18	.18	1.000	n.s.
Group III	2.45	.394				
Group II	2.11	.586	–.34	.19	1.789	n.s.
Group III	2.45	.394				
One Hour						
Group I	.61	.726	–.23	.29	.793	n.s.
Group II	.84	.842				
Group I	.61	.726	–.85	.33	2.576	.05
Group III	1.46	.991				
Group II	.84	.842	–.62	.35	1.771	n.s.
Group III	1.46	.991				
72 Hours						
Group I	.27	.384	–.36	.22	1.636	n.s.
Group II	.63	.726				
Group I	.27	.384	–.81	.28	2.893	.05
Group III	1.08	.964				
Group II	.63	.726	–.45	.32	1.406	n.s.
Group III	1.08	.964				

ing the responses and the experimenter were adequate for the purpose of this study.

Analysis and Results

In the analysis of data, 45 responses in each production trial for each subject were combined in one distribution. This procedure produced a mean response score for each individual for each production trial and indicated the quality of his response to all three of the sounds, rather than to each sound separately.

Table 1 shows the trends in the mean response scores for Groups I, II and III in the three test situations. An examination of the data presented in this table indicates that Groups I, II and III did not vary greatly in the immediate test situation. (It will be noted later in Table 2 that, using a *t*-test from Lewis (2), the differences between Groups I, II and III were not significant at the five per cent level and therefore the groups were assumed to be homogeneous). This might have been expected since the three groups had received the same number of stimulation cycles before the immediate test situation. On the other hand, Table 1 indicates that the greater the amount of training, the better the retention of the sound. This was shown by the increase of the mean scores from Group I through III in the second and third testing

TABLE 3. The means of, and the difference between the means of the immediate, the one hour and the 72 hour test situations for groups I, II and III. (N=15)

	Mean	*S.D.*	*D*	*S.E. diff.*	*t*	*Sig. level*
Group I						
Immediate One Hour	2.27 .61	.565 .726	1.66	.18	9.222	.01
Immediate 72 Hours	2.27 .27	.565 .384	1.99	.14	14.214	.01
One Hour 72 Hours	.61 .27	.726 .384	.34	.16	2.125	n.s.
Group II						
Immediate One Hour	2.11 .84	.586 .842	1.26	.16	7.875	.01
Immediate 72 Hours	2.11 .63	.586 .726	1.48	.15	9.867	.01
One Hour 72 Hours	.84 .63	.842 .726	.14	.11	1.273	n.s.
Group III						
Immediate One Hour	2.45 1.46	.394 .991	.99	.27	3.667	.01
Immediate 72 Hours	2.45 1.08	.394 .964	1.37	.25	5.480	.01
One Hour 72 Hours	1.46 1.08	.991 .964	.36	.22	1.636	n.s.

situations. This might be expected since Group III received more stimulation cycles than Group II and Group II received more than Group I.

In Table 2, comparisons of the means and the differences between the means were made between Groups I and II, Groups I and III, and Groups II and III in the immediate, the one hour and the 72 hour test situations. This comparison was made to determine the effect of the additional stimulation upon the production of the sounds.

Table 2 indicates that Group III differed significantly from Group I at both the one hour and the 72 hour test situations. However, Group I did not differ significantly from Group II nor did Group II differ significantly from Group III in either test situation.

This table shows that the ability to produce a sound after a time lapse is affected by the amount of stimulation given. A small amount of additional stimulation produced no significant gains, but larger amounts of stimulation did.

Table 3 shows the differences between the means of the immediate and the one hour test situation, the immediate and the 72 hour test situation, and the one hour and the 72 hour test situation for Groups I, II and III. A *t*-test for related data from Lewis (*2*) was used.

This table indicates that the ability to reproduce a sound is proportional to the length of time elapsing since the subject was stimulated. The lapse of time immediately after stimulation produced a greater loss of production ability than any proportional period of time later on. Thus a large and significant loss was observed between the immediate and the one hour test situation. A smaller but not significant loss was shown between the one hour and 72 hour test situations. Of course, the immediate and 72 hour test situations differed significantly.

Summary

The purpose of this study was to investigate the effect of increased training, involving stimulation and reinforcement, upon an individual's skill in production and retention of unfamiliar sounds. Information of this type is needed in articulation therapy to determine the reliability of the concept that continued stimulation of a sound will bring about better retention.

Three groups of high school freshmen were used in this study. Group I received only a minimal amount of stimulation, that is, five stimulation cycles. Group II received five additional stimulation cycles. Group III received 15 additional stimulation cycles.

Each subject gave 15 responses to each of the three sounds used in this study, during each of the three testing situations: i.e., the immediate, the one hour and the 72 hour test situations.

Conclusions

1. The larger amounts of stimulation resulted in significantly better ability to reproduce sounds after a time lapse. Although smaller amounts of stimulation did produce tendencies toward improvement in retention, these trends were not statistically significant.

2. The ability to reproduce a sound was affected by the time lapse between the training of the subject and his production of the sound. A time lapse immediately after training resulted in the greatest loss in ability to produce the sound. This loss in

ability was partially counteracted by increasing the amount of training.

3. The amount of stimulation of any sound for any subject was comparatively small. For example, the subjects in Group III received the greatest amount of stimulation and yet they were given only about two minutes of stimulation per sound. This is an extremely small amount of stimulation when compared with the amount given in a clinical situation. In general, the clinician might anticipate an increase in retention following increased stimulation. A clinician must, however, expect large initial losses in retention shortly after the clinical session is completed and should not be discouraged by these losses. It might be wise, however, not to rush into complicated speech patterns until the more simple ones are well established. The clinician might obtain a 'Basic retention measure' by testing a few days after the clinical session.

References

1. JOHNSON, W., S. F. BROWN, J. F. CURTIS, C. W. EDNEY AND J. KEASTER. *Speech Handicapped School Children.* New York: Harper, 1948.
2. LEWIS, D. *Quantitative Methods in Psychology.* Iowa City: Gordon Bookshop, 1948.
3. LINDQUIST, E. F. *A First Course in Statistics.* New York: Houghton Mifflin, 1942.
4. TRAVIS, L. E. *Speech Pathology.* New York: Appleton-Century, 1931.
5. VAN RIPER, C. *Speech Correction: Principles and Methods.* (2nd ed.) New York: Prentice-Hall, 1947.