The Relationship Between Remedial Speech Instruction Activities and Articulation Change

LINDA CHISUM, M.A.
RALPH L. SHELTON, Jr., Ph.D.
WILLIAM B. ARNDT, Jr., Ph.D.
MARY ELBERT, M.A.
Kansas City, Kansas

Acquisition of motor skills is thought to be facilitated by practice (*3*, *5*) and by reinforcement of correct responses (*6*). Both principles are applied in the remediation of defective articulation, but the relationship between activities used in speech remediation and measures of gain in articulation skill has not been examined. The purpose of this investigation was to study the relationship between change in articulation score before and after a period of instruction and two independent variables: practice responses by the learner and reward responses by the instructor. Examination of this relationship requires evidence that the learners improved their articulation in response to the treatment provided. Therefore, change in articulation score over a period during which no treatment was provided was compared with change during a treatment period comparable in length to the control period. The findings are pertinent to the understanding of the remedial development of articulation in both palate defective individuals and persons with other kinds of articulation defects.

Procedure

SUBJECTS. Originally, 17 subjects were to be tested before and after a six-month nontreatment period and again after a six-month period of speech instruction. However, some of the subjects who qualified for the control period did not complete the minimal number of lessons. Additional subjects were therefore studied during an instructional period only. Thus control data were obtained on twelve subjects and experimental data on eleven. These will be referred to as control and experimental groups and only six of the 12 and 11 persons served in both groups.

Mrs. Chisum, Dr. Shelton, and Mrs. Elbert are with the Hearing and Speech Department of the University of Kansas Medical Center. Dr. Arndt is with the Psychology Department, University of Missouri at Kansas City.
Some of the findings in this study were reported at the 1966 meeting of the American Speech and Hearing Association in Washington, D. C. This study was part of a project supported by PHS Research Grant DE-02004-03, National Institute of Dental Research.

The control group included nine persons with cleft palate and three with palatal insufficiency. Seven of the cleft palate subjects had surgically repaired palates; the others wore prosthetic appliances. One of the surgically repaired subjects had a pharyngeal flap. The experimental group consisted of eight subjects with cleft palate and three with palatal insufficiency. Six of the subjects with clefts had had surgical repair and two wore speech appliances. The palatal insufficiency was identified on the basis of the subjects' articulation and voice characteristics. All subjects had either hypernasality or consonant distortion by nasal escape of air, or both. Although their speech was sufficiently disordered to motivate the subjects' parents to seek professional help, the subjects were considered to have borderline palatopharyngeal deficiency and mild-to-moderate speech problems. For inclusion, each subject had to have an articulation problem involving a minimum of three phonemes and a total of at least six errors on a 223-item articulation test. The articulation test used permitted study of articulation gain in terms of both change in total test score and change in different phonetic categories including plosive, sibilant, fricative, glide, and nasal phonemes.

The subjects, who ranged in age from six to twelve years, were required to obtain minimum scores on the Peabody Picture Vocabulary Test, the Auditory-Vocal Sequencing Subtest of the Illinois Test of Psycholinguistic Abilities, and the Boston University Speech Sound Discrimination Picture Test. In addition, a pure-tone threshold audiometric test was administered so that persons with a hearing loss greater than 20 dB in the better ear (ASA) did not participate in the study. Thus an attempt was made to eliminate individuals with gross problems of intelligence, auditory perception, or hearing acuity which might adversely influence their articulation learning.

TEACHING DESIGN. Each experimental subject received a 30-minute individual speech lesson twice weekly for a minimum of 30 lessons within a six-month period or for a maximum of 48 lessons within a nine-month period. The mean instruction period was 7.18 months; the mean control period was 6.83 months.

The method of teaching involved a motor learning approach with emphasis on response-shaping procedures. If the subject was unable to produce a sound in isolation, he was instructed to attempt to produce it with the articulators in various positions. Successive approximations of the correct response were rewarded by the clinician as the subject learned the desired response. The instructor also advised the subject to watch and to listen while the instructor produced the desired response. Having presented a verbal model of the response, the instructor asked the subject to imitate the response. The instructor then reported to the subject whether the response was acceptable. If the response was not correctly produced, the instructor made suggestions which he considered beneficial in helping the subject accomplish the desired response.

Some subjects first learned the correct response in isolation from context and then progressed through nonsense syllables, words, sentences, and conversation. However, if a subject was unable to learn a phoneme in isolation after repeated trials, an attempt was made to teach it in various contexts. If a subject could produce his goal phoneme at the beginning of the instruction, work was directed to use of the phoneme in words and in longer contexts. The instructors were careful to avoid tasks on which a subject experienced repeated failure. All speech practice was restricted to the speech instruction sessions. No assistance from parents was solicited. Four persons provided the instruction, but no subject was taught by more than one instructor.

ANALYSIS OF INSTRUCTION TAPES. Speech instruction sessions were tape recorded so that certain counts could be made from the tapes and examined for correlation with measures of articulation change. Two tape recordings were selected randomly from those collected during each quarter of the instruction period. These were evaluated by an experimenter who was instructed to count the practice responses of the subject and the reward responses of the instructor. Three categories of practice responses were tabulated: practice of phonemes, practice of words, and practice of sentences or longer units. Practice responses included a subject's first imitation of a stimulus presented by his instructor and immediate additional repetitions made even though the instructor did not repeat the stimulus.

Instructors' reward responses were categorized as specific reward or general. A reward was specific if given immediately after the subject's response or if the clinician verbally related the reward to a particular response. Specific rewards included verbal comments and tokens. General rewards were not related to specific articulatory responses but involved praising the subject for doing well during a series of responses or during an entire session.

JUDGE RELIABILITY. Four speech clinicians participated in articulation test administration. These were the same persons who provided the therapy. However, no clinician administered a post-instruction articulation test to a subject whom he had taught. Persons administering the articulation tests were generally unaware of subject's status in the study and never knew scores from previous tests. Articulation responses were scored at the time they were made and were not tape recorded.

Before administering tests to subjects, three of the four clinicians scored live articulation test responses from children with articulation problems. The fourth clinician scored articulation responses from tape recordings which had been scored previously by two other clinicians. Pearson correlation coefficients, comparing total scores assigned by judge pairs, were .98, .99, and .99, for live judgments and .86, .93, and .93, for the taped data. Percentages of agreement, comparing judge pair responses to articulation test items, were .86, .88, and .90, for live data and .77, .83, and .87, for data from tapes. Overall articulation judge agreement was considered to be satisfactory. In later analysis of data for three subjects, disagreement on the presence and absence of consonant distortion by nasal escape of air may have existed. This will be discussed below.

Two judges participated in the counting and the classification of responses from the instruction tapes. Correlation coefficients for intra- and inter-judge response counts for the various categories of practice and reward responses ranged from .82 to .99, with one exception: the category of general reward responses. One intra-judge coefficient was only .42, and the Pearson coefficient between the two judges was .43. Thus judge agreement was considered to be satisfactory, except for the one category.

Results

ARTICULATION CHANGE. Control group members were administered two articulation tests as were experimental group members. Means and standard deviations for total error score on these test administrations are reported in Table 1. The difference between each subject's scores on the two tests was computed, and means and standard deviations for these difference scores are also reported in Table 1. Difference scores indicate the amount and direction of change in articulation test performance between tests.

Several t tests, also included in Table 1, were used to study change in articulation score in the control group and in the experimental group and to compare the control group difference score with the experimental group difference score. The difference between the scores for the first and second test was not significant for the control group (t, 1.55), but was significant for the experimental group (t, 4.22). (The t for matched groups was used for these analyses.) Thus the experimental group significantly reduced its articulation errors while the control group did not. A third t test, this time for randomized groups, was computed to compare the means of the two sets of *difference* scores. The obtained t ratio of 2.73, significant at the 5% level, indicated a greater reduction in total test articulation errors by experimental than by control subjects.

TABLE 1. Means and standard deviations for the articulation tests given to the control and experimental subjects and for differences between the two sets of scores within each subject group. Data are for scores on the total test. t ratios comparing articulation test scores and articulation test difference scores are also included. Asterisked values are significant at the 5% level.

	Control				*Experimental*				
	Pre-score	*Post-score*	*t*	*Difference score*	*Pre-score*	*Post-score*	*t*	*Difference score*	*t for difference score*
M	52.33	45.00	1.55	7.33	64.81	33.36	4.22*	31.45	2.73*
SD	32.98	25.47		16.38	12.94	18.59		23.52	

TABLE 2. Experimental group articulation test score means and standard deviations for five phonetic categories. t ratios comparing the sets of scores are also reported. Asterisked values are significant at the 1% level.

	Pre-scores		*Post-scores*		
	M	SD	M	SD	t
Plosives	14.18	12.80	5.55	5.41	2.12
Sibilants	39.36	13.86	21.18	13.01	4.77*
Glides	3.55	3.88	1.36	1.63	2.21
Fricatives	11.63	9.31	4.73	5.76	3.94*
Nasals	.36	.81	.27	.64	.32

TABLE 3. Pearson correlation coefficients for articulation difference scores and measures of practice and reinforcement. None of the coefficients was significant at the 5% level. Coefficients between reward and sibilants and fricatives were not computed because of the low correlations for combined sibilants and fricatives.

Test	*Practice*			*Reward*		
	Sounds	*Words*	*Sentences*	*Total*	*General*	*Specific*
Total	.47	.14	−.01	.46	−.10	.23
Sibilants, fricatives, combined	.39	.09	−.03	.37	−.19	−.15
Sibilants	.22	.04	−.05	.20	—	—
Fricatives	.54	.14	.004	.52	—	—

Next, the number of errors which experimental group members made on plosives /p, b, t, d, k, g/, sibilants /s, z, ʃ, ʒ, tʃ, dʒ/, glides /r, l, j/, fricatives /f, v, θ, ð/ and nasals /m, n, ŋ/, was determined for both the pre- and post-instruction articulation tests.[1] Means, standard deviations, and *t*s are reported in Table 2 for each phonetic category. The results indicate that the number of sibilant and fricative errors was significantly reduced. The persons who provided the instruction reported that they emphasized sibilants in their work.

CORRELATION COEFFICIENTS BETWEEN ARTICULATION CHANGE AND RESPONSE COUNTS. The articulation difference scores for experimental subjects were correlated with measures of practice and reinforcement. Specifically, as shown in Table 3, difference scores for the total articulation test, sibilant items, fricative items, and sibilant and fricative items com-

[1] The original study plan called for comparison of articulation errors on words produced spontaneously in response to pictures, on imitated words out of context, and on imitated words in sentences. The number of errors on imitated words in sentences was significantly greater than the number of errors on words out of context. However, no other differences were statistically significant, and the decision was made not to pursue these test category influences further.

bined, were each correlated with measures of subject practice responses (practice sounds, words, sentences or longer units, and total practice responses) and clinician reward responses (general and specific). The subcategory of total practice responses included the total of practiced sounds, words, and sentences or longer units. None of the coefficients was significant at the 5% level, indicating that there was no statistically significant relationship between articulation gain and any of the practice and reinforcement response counts. Possible reasons for the low correlations are presented below.

Discussion

CONSONANT DISTORTION BY NASAL ESCAPE OF AIR. As reported, reliability of judges' assessment of articulation errors on a right-wrong basis for the experimental and control groups was adequate except for disagreement concerning three subjects. Since the number of subjects involved was small, any articulation test-scoring error reflected by the disagreement probably had little effect on the results. The judges had typically identified the nature of articulation errors after deciding whether the response was correct or incorrect. From examination of the test protocols and from reexamination of one of the subjects, it was learned that one judge disagreed with others in observation of consonant distortion by nasal escape of air. These deliberations indicate that a distinction should be drawn between consonant distortion when the air stream is directed out the nose and sufficient oral pressure for the production of the phoneme does not exist, and consonant distortion, when the sound is well articulated orally but is accompanied by noise in the nasal passages presumably caused by slight leakage of air past the elevated palate or the obturator. It is thought that the latter kind of sound production was marked as erroneous production by one judge but not by others.

RELATIONSHIP BETWEEN ARTICULATION CHANGE AND RESPONSE COUNTS. The importance of practice and reinforcement to motor learning is so well established that other factors must be found to account for the relatively low correlations obtained in this study. Poor measurement reliability would account only for the low correlation between articulation gain and general reward responses. One possible explanation is that the measurements of articulation or of subject and clinician responses were not valid. The validity of articulation testing procedures involving the single consonants of English has been demonstrated by correlation procedure wherein test scores were correlated with judgments of subjects' articulation adequacy (*4*). Nevertheless, it is possible that the counts of incorrect phoneme production in the contexts used in this study did not reflect accurately the articulation improvement made by the subjects. If articulation acquisition is viewed as involving an acquisition phase and an automation phase, it is possible that a subject could show error-free articulation test behavior while still requiring help in automating the newly-acquired responses.

Additional instruction would increase the practice and reward counts without changing the articulation scores. Thus if the articulation testing procedure used is valid only for measurement of change during the acquisition phase, then termination of data gathering near articulation acquisition asymptote might have resulted in higher correlations.

The definitions used in counting practice and reward responses were described by Chisum (*1*); however, the validity of the criteria for defining response counts has not been established. Use of devices to count responses of interest at the time of the instruction would avoid possible problems associated with tape recording responses where noise and mouth-to-microphone distance are not well controlled. This would also eliminate problems resulting from sampling the obtained tapes.

The accuracy of the instructors in rewarding correct responses or suitable approximations of the correct responses is another possible source of measurement error. Variability from lesson to lesson in amount of practice or in reinforcement of faulty responses would influence the correlations obtained. In future studies which involve face-to-face evaluation, an effort should be made to observe and control those sources of variability. The advantages and disadvantages of controlling the schedule of reinforcement and the number of practice responses should also be considered.

Another possible explanation for the low correlations involves the relationship between rate of learning and number of practice responses. For instance, one individual may require 100 practice responses to change an incorrect response to a correct one, whereas another individual may require 200 practice responses to accomplish the same task. The relationship between rate of learning and treatment activities should be studied in future investigations. In addition to testing subjects at the beginning and at the end of a lesson series, articulation tasks could be constructed involving sounds to be taught or observed (*2*).

Summary

The purpose of this investigation was to study the relationship between change in articulation test score where articulation is measured before and after a period of instruction and two independent variables: practice responses by the subject and reward responses by the clinician. Evidence is also given of the subjects' articulation improvement in response to the instruction provided. Judge reliability for the task of scoring the articulation test responses as correct or incorrect and for the task of counting practice and reward responses from tape-recorded speech instruction sessions was found to be satisfactory. The result of t tests indicated that as a group the experimental subjects did reduce the number of articulation errors, presumably as a result of instruction. Additional statistical tests showed that significant gains were made in the sibilant and fricative sound categories. Pearson correlation coefficients between articulation gain and subject and clinician response counts were computed and were not signifi-

cant. Possible explanations for the low correlations and procedures for use in future studies are presented in the paper.

reprints: *Dr. Ralph L. Shelton*
Hearing & Speech Department
University of Kansas Medical Center
Kansas City, Kansas 66103

Acknowledgments. The authors wish to thank Drs. Arthur F. Lindquist, Frank W. Masters, and David W. Robinson, of the Kansas University Medical Center Cleft Palate Team for their cooperation in securing subjects and Mrs. Delores Crary for participating as an instructor in the study.

References

1. Chisum, Linda, Relationship between remedial speech instruction activities and articulation change. Unpublished Masters Thesis, University of Kansas, 1966.
2. Elbert, Mary, R. L. Shelton, Jr., and W. B. Arndt, Jr., Development of a task for evaluation of articulation change. *J. speech hearing Res., 10,* 281–288, 1967.
3. Fitts, P. M., Factors in complex skill training. *Training, Research and Education,* Robert Glaser, ed., 177–197. New York: John Wiley and Sons, Inc., 1962.
4. Jordan, E. P., Articulation test measures and listener ratings of articulation defectiveness. *J. speech hearing Res., 3,* 303–319, 1960.
5. Paillard, J., The patterning of skilled movements. *Handbook of Physiology,* J. Field and others (eds), Vol. III, 1679–1708. Washington, D. C.: The American Physiological Society, 1960.
6. Kimble, G. A., *Conditioning and Learning,* Second Edition. New York: Appleton-Century-Crofts, Inc., 1961.