JOURNAL OF APPLIED BEHAVIOR ANALYSIS 1977, 10, 717-733 NUMBER 4 (WINTER) 1977

GENERALIZING ARTICULATION TRAINING WITH TRAINABLE MENTALLY RETARDED SUBJECTS[1]

JANE Y. MURDOCK, EUGENE E. GARCIA, AND MICHAEL L. HARDMAN

UNIVERSITY OF NEW ORLEANS, UNIVERSITY OF UTAH, AND UNIVERSITY OF UTAH

A multiple-baseline technique was used to evaluate generalization effects during articulation training with trainable mentally retarded subjects. Four target words were selected for each subject on the basis of whether the subject could articulate the word correctly when it was modelled but could not articulate the word correctly in response to a picture of it. Five different settings were selected for generalization probing and training for each subject. In Setting 1, Experimenter 1 initiated training sequentially on all four target words for each subject. Other experimenters probed for correct articulation generalization in four other settings. Training was initiated in these four other settings sequentially only if correct responding failed to generalize to a setting. Results indicated that it was necessary to initiate training on at least three of the four selected target words in at least one additional setting with an additional trainer before correct responding generalized to untrained settings.

DESCRIPTORS: articulation, generalization, retardates

The failure to generalize is one of the most commonly noted characteristics of the retarded person. This is particularly so with the retarded individual who is classified at a level of trainable or below. The purpose of this study was to attempt to determine if a relatively simple generalization training program would result in correct articulation generalization with trainable mentally retarded subjects.

Speech and language generalization requires that the individual be able to respond correctly on all language tasks that have been mastered with the original language instructor (1) in other physical settings, (2) with other people, (3) to different discriminative stimuli and, (4) to similar language tasks. One of the most difficult tasks has been establishing generalization of correct articulation responses (Mowrer, 1971; Wing and Heimgartner, 1973). Typically, speech correctionists report that the training has been successful and need not continue, but teachers and parents complain that the child's articulation has not improved. Obviously, the correct articulatory responses have not generalized to other people in other settings outside the confines of the training situation. There is frequently some mention in speech-training programs regarding training in other settings, calling for other people to engage in training, usually parents, teachers, or other personnel in the particular institution (Risley and Wolf, 1967; Sloane, Johnston, and Harris, 1968; Van Riper, 1963; Wing and Heimgartner, 1973). Unfortunately, these generalization procedures usually are not clearly specified and/or there is no systematic documentation of whether or not generalization occurred. The following literature is related to articulation programs and investigations that have clearly specified their generalization procedures and that have specifically documented the extent to which generalization occurred.

[1]An earlier draft of this manuscript was submitted by the senior author to the Department of Educational Psychology, University of Utah, in partial fulfillment of the requirements for the PhD degree. The authors wish to thank all of the teachers at Garfield School in the Salt Lake City School District who participated in conducting this study as well as the Principal, Gene Smith, and other administrators for their cooperation. Reprints may be obtained from Jane Y. Murdock, Department of Special Education, ED 246, University of New Orleans, New Orleans, Louisiana 70122.

Mowrer, Baker, and Schutz (1968) devised an instructional training manual for parents of lisping children. The training manual explained precisely what the parent was to do, say, and evaluate while training the child at home. The authors reported that when parents followed the instructions explicitly, 70% of the children articulated the /s/ phoneme correctly for the first 20 /s/ responses in a test of conversational speech.

Ryan (1971) made a home or transfer procedure part of his articulation training program for children who lisped. Parents, teacher-aides, volunteer adults, older children in the school, another child in the home, or a sibling were trained to discriminate the correct articulation of the target sounds, and the time, place and reinforcers to be used at the transfer site. On the basis of a 2-min conversational speech sample, the mean percentage of correct articulation for the trained phonemes was 77.11% with a range of 17% to 100%.

Costello and Bosler (1976) also designed a program in which the parents did the articulation training in the home with three children who had functional articulation problems. When the children reached the success criterion on the /v/ phoneme in the home, generalization was probed in four additional settings. Costello and Bosler found that correct articulation generalized to nontreatment settings.

Raymore and McLean (1972) used a procedure based on the application of operant principles to gain stimulus control over a mildly retarded child's newly learned phoneme responses. Their subject had an IQ of 69. Their concern was more with the form of evoking the response and also with generalization of new phoneme responses into initial, medial, and final positions within words. They felt that training the phoneme in all three positions precluded the need to train for generalization across settings or with other people.

There is considerable literature relating to establishing correct grammatical forms with moderately and severely retarded individuals (Baer and Guess, 1971; Garcia, 1974; Garcia, Guess, and Byrnes, 1973; Gray and Ryan, 1973; Guess, Sailor, Rutherford, and Baer, 1973; Sailor, 1971; Schumaker and Sherman, 1970). Garcia (1974) investigated generalization of a conversational speech form and found that profoundly retarded subjects did not generalize to someone who had not participated in the training until at least two other persons had trained the response.

There seems to be a dearth of literature reporting on either training or generalization of articulation skills with the moderately or severely retarded child. It is possible that there is still a feeling that articulation training is hopeless with retarded individuals (West, Ansberry, and Carr, 1957). It is also possible that speech correctionists and investigators do not consider articulatory errors as being serious enough to warrant training in the face of the multiple language handicaps that are usually present in the mentally retarded person. However, it would seem that the more limited a person's communicative repertoire is, the more critical it is that he or she communicate whatever possible as intelligibly as possible. The present study examined procedures that would facilitate generalization of correct articulation responses of trainable mentally retarded children to people who had not participated in the original training and to nontraining settings. Snyder, Lovitt, and Smith (1975) reviewed language training of the severely retarded and noted that in every study they reviewed, some form of tangible reinforcer was used as part of the training package. Thus, only social consequences were delivered immediately following a correct response in the present study.

METHOD

Design

A multiple-baseline design was used to evaluate the effects of the experimental procedures (Baer, Wolf, and Risley, 1968). This required ongoing measures of four different ver-

bal responses for each subject in each of five different settings. After baseline in all five settings, experimental procedures were initiated in Setting 1 for the first target word. Baseline procedures were maintained for the remaining three target words in Setting 1 and for all target words in Settings 2, 3, 4, and 5. As each subject met the predetermined criterion for success on the first target word in Setting 1, the experimental procedures were initiated for that word in Setting 2, unless the subject was already responding correctly in that setting at a criterion level of 75% or better for two consecutive sessions. These procedures were continued until the first target word was either being trained in four of the five settings or the subject had generalized correct responding at 75% or better without any training. Experimental procedures were never initiated in Setting 5, thus providing a possibility for each subject to generalize to an untrained setting throughout the study. These procedures were systematically replicated with each subject for the three remaining target words.

Subjects

Subjects were selected from two classrooms in a public school for the trainable mentally retarded in Salt Lake City, those selected could not articulate the target words correctly in response to a picture representing the word, but could articulate the target word correctly when the experimenter modelled the correct articulation. Subject 1 was 8 yr 11 months at the beginning of this study. Her reported IQ from a Stanford-Binet was 44. The cause of her retardation was unknown. She was given the Photo Articulation Test (PAT) and made 22 articulation errors when a single-word response to a photograph of the object was required. Most of these errors were to frequently occurring sounds in English usage, including /g/, /z/ and all other sibilants, /r/, /th/, and all of the blends included in the test. For the /g/ sound, she substituted /d/ in the initial and medial positions and omitted the sound in the final position. For the /z/ sound, she substituted a distorted /s/ in the initial position, a /d/ in the medial position, and omitted the sound in the final position. She could articulate all of these correctly immediately after listening to and watching a model. Consequently, the words *zebra, hose, gun,* and *tiger* were selected as target words.

Subject 2 was 9 yr seven months at the beginning of this study. She had a medical diagnosis of Down's syndrome. Her reported IQ from a Stanford-Binet was 44. She was given the PAT and made 32 articulation errors when a single-word response to a photograph of the object was required. Most of these errors were to frequently occurring sounds in English usage, including /v/, /r/, /s/ and all other sibilants, /th/, and all of the blends included in the test. For the /v/ sound she had a severe distortion in the initial position, a /b/ substitution in the medial position, and an /f/ substitution in the final position. For the /r/ she substituted /w/ in the initial and medial positions, and omitted the sound in the final position. She could articulate these sounds correctly immediately after listening to and watching a model. Consequently, the words *vacuum, TV, radio,* and *doctor* were selected as target words.

Materials

The materials consisted of stimulus pictures selected from the Photo Articulation Test and the Peabody Language Development Kit, Level 1, which included pictures of each subject's four target words: a zebra, a hose, a gun, and a tiger for Subject 1 and a stove, a TV, a radio, and a doctor for Subject 2. "Articulation Training and Generalization" cards were designed for each subject. These cards clearly specified criteria for correct articulation responses, settings, probes, spaces for recording a correct (+) or incorrect (—) response to each probe, the reinforcement schedule or a notation that no reinforcement was to occur, and explicit instructions regarding training procedures whenever such training was to occur. Tape recorders

were used to record all articulation probing and training trials. Videotape recordings of sessions were made at least once for each subject during each phase of the study in every setting.

Settings and Experimenters

Settings and experimenters varied for each subject. However, for all subjects, Setting 1 was with E-1 (a female) in a small 2.4 by 3 m room. The room contained books, materials, and usual office furniture. Each subject was seen individually by E-1 in this room.

For Subject 1, Setting 2 was with E-2 (a female) in a corner of the regular classroom. It was a special language-development session designed for students who did not have the language skills of a normal 4-yr-old (Murdock and Hartmann, 1975). One other student was present. Setting 3 was with E-3 (a male) in a learning center setting where individualized academic activities varied for each of the 13 students in the center. Setting 4 was with E-2 again in a Peabody Language, Level 1, session, which was conducted in the hallway. Four other students were present. Setting 5 was with E-4 (a female) in a regular classroom lunch period with 20 students present. E-4 did not train Subject 1 in any setting.

For Subject 2, Setting 2 was with E-5 (a female) in a corner of a regular classroom. It was a math session with six other students present. Setting 3 was with E-6 (a female) in a corner of a different classroom. It was a reading session with six other students present. Setting 4 was also with E-6 in another corner of the same classroom. It was a Distar Language 1 session with six other students present. Setting 5 was with E-5 in a regular classroom lunch period with 20 students present.

Every effort was made to vary both the settings and the experimenters as much as possible within the context of a public-school situation, where both were at a premium. For Subject 1, the only physical settings that were the same were Settings 2 and 5, and the experimenter was the same in Settings 2 and 4. For Subject 2, the physical settings and the experimenters were the same in Settings 2 and 5 and 3 and 4, respectively. However, for both subjects the other students present varied in all settings, as well as the literal location of the particular settings within the same classroom. It was hoped that 20 children seated at tables during lunch would constitute a considerably different environment than six students seated in a semicircle in a corner of the room.

Procedures

Probe trials. Probe trials consisted of presenting each subject with a stimulus picture and asking, "What is this?" four times for each target word (a total of 16 probes for Settings 2, 3, 4, and 5). A total of 48 probes, or 12 per target word were presented in Setting 1. A larger number of probes in Setting 1 was used to expedite training when it was initiated. Only a single-word response was required, but any appropriate chain that included the target word was acceptable. No modelling or reinforcing occurred following the subjects' responses during probing. Generalization probing was continued in all settings for all target words not being trained in each particular setting.

Training trials. Training consisted of immediate social consequences for a correct response or immediate modelling of the correct response following an incorrect response. All training was on a continuous reinforcement (CRF) schedule; thus, every correct response was socially consequated. At the end of a training session, Subject 1 received points and/or candy and Subject 2 received candy. These reinforcement procedures were consistent with those used in each subject's classroom.

Phases of the study. Baseline records were taken for each subject in each of the five settings through the probing procedures. When a subject had three or more consecutive sessions with 25% or less correct responding for all four target words in all five settings, training was initiated in Setting 1 for Target 1. The target words trained in each setting during the suc-

cessive phases of the study varied from subject to subject, depending on whether or not the subject had generalized to untrained settings by responding correctly at the criterion level of 75% or better for two consecutive sessions.

Preliminary training for experimenters. Since the immediate consequences used in this study were social consequences, it was necessary to set up certain procedural controls that may not have been required if tangible rewards had been used. Social consequences can be extremely subtle. A glance, a wink, or a touch may be sufficient to maintain or strengthen behaviors even when these are not accompanied by any verbal praise. To control for this, experimenters who participated in this study were instructed not to look at, smile at, touch, or verbally praise any subject following correct or incorrect responses made during baseline or probe trials.

To determine whether the experimenters concurred as to what constituted a correct response, auditory discrimination training was conducted before any of the experimenters trained any subject's articulation. E-1 presented tapes of the subjects after training had been initiated in Setting 1. The tapes contained mixed correct and incorrect responses. These tapes were listened to, scored, and discussed by each experimenter involved, including E-4 who conducted only generalization probing throughout the study. The training was continued until a minimum of 90% agreement was reached between E-1 and all other experimenters listening to the same subject for each of the target words. This auditory discrimination training was evaluated daily by E-1, who listened to and scored the tape recordings of all probing and training that occurred in all settings.

Scoring and Reliability

All training and probing trials were tape recorded in their entirety. Each experimenter scored the subjects' responses as correct or incorrect. (There were no incidents of either subject not responding.) These procedures were followed so that scoring was based on auditory cues alone, rather than on both auditory and visual cues. E-1 scored the tape-recorded responses from all other settings daily, both for purposes of auditory discrimination training and reliability. The reliability of E-1's scoring was checked by E-3, who scored randomly selected portions of tape recordings from Setting 1. This was done for a minimum of 20% of all training and probing done by E-1 and was also done at least once during each phase. The resulting data were cast into 2 x 2 contingency tables for purposes of reliability measurement. The percentage agreement was calculated by dividing the total number of agreements by the total number of agreements plus disagreements and multiplying by 100. The overall percentage agreement ranged from 92% to 98% for both training and probing. A phi coefficient, a percentage agreement value that corrects for expected agreements, was also then applied to the 2 x 2 table data (Hartmann, 1977). The overall phi coefficient ranged from 0.78 to 0.97. For this reason, not only was an overall phi coefficient calculated for each target word, but one was also calculated for each target word during a phase or combination of phases where the number of correct and incorrect responses was as nearly equal as possible. These phi coefficients ranged from 0.59 to 0.96, and the percentage agreement for these same phases ranged from 83% to 98%.

To determine whether the daily auditory discrimination training resulted in the desired accurate auditory discrimination, rather than consensual observer drift (Johnson and Bolstad, 1973), a tape recording was made at the end of the study of sections from each of five different phases. Each section included 80 trials or a total of 400 trials for each experimenter. These sections were recorded randomly, and a key prepared by E-1 to designate which phase and which setting each section of the tape had been recorded from. This tape was then given to the experimenter who had originally scored these sections to get intraobserver reliability, *i.e.,* the scoring of these 400 trials at the end

of the study was compared to the same 400 trials scored by the same experimenter while the study was in progress. In this way it could be determined whether the experimenters maintained consistent criteria for correct responding throughout the study or whether consensual observer drift had occurred. A phi coefficient and a percentage agreement were calculated to determine this intraobserver reliability, resulting in 0.89 and 94% respectively.

Procedural reliability was determined by videotaping at least once for each experimenter during each phase of the study. Videotaping was a common occurrence in these classrooms. Consequently, the experimenters would have no reason to associate this videtotaping with this study. E-1 and E-3 simultaneously but independently judged whether the observed experimenters smiled, had physical contact, had eye contact, or verbally praised a subject immediately following the subject's responses during probing. They also judged whether the experimenters appropriately reinforced or modelled during training. The percentage agreement between the scores obtained by E-1 and E-3 was calculated by dividing the total number of agreements by the total number of agreements plus disagreements and multiplying by 100. The percentage agreement was 99.5%.

RESULTS

Training and Generalization: Subject 1

Training. In Setting 1, Subject 1 required seven training sessions to reach criterion on Target 1 (see Figure 1), two training sessions to reach criterion on Target 2 (see Figure 2), three sessions to reach criterion on Target 3 (see Figure 3), and six sessions to reach criterion on Target 4 (see Figure 4). The criterion was correct responding at a 75% level or better for two consecutive sessions. Consequently, a subject could not reach criterion in less than two sessions.

Generalization. During baseline, there was no correct responding for Targets 1, 2, and 3 in any setting (see Figures 1, 2, and 3). There was erratic correct responding up to 50% during Session 9 for Target 4 in Settings 2 and 3. For the last nine sessions of baseline, correct responding for Target 4 was 0% in all settings. (See Figure 4.)

After training was initiated during Phase 1 in Setting 1, there was unstable correct responding for all four targets in all four settings, ranging from 0% to 100%. Correct responding did not stabilize or reach criterion for any target in Settings 2, 3, 4, and 5. Consequently, training was initiated in Phase 2 in Setting 2, after which correct responding for Targets 2 and 4 reached criterion and stabilized in Settings 3, 4, and 5. Subject 1 responded on Target 2 at 75% or better for the last 19 sessions of the study and for Target 4 at 100% for the last 15 sessions of the study. (See Figures 2 and 4.)

However, after training was initiated during Phase 2, correct responding for Targets 1 and 3 continued to be unstable and did not reach criterion. Consequently, training was initiated in Setting 3 during Phase 3, after which correct responding reached criterion and stabilized in Settings 4 and 5. The subject responded correctly at 100% on Target 1 for the last 19 sessions of the study (see Figure 1). Phase 3 was not initiated for Target 3 until Session 45 because of the multiple-baseline requirements of the study. However, the subject responded correctly at 100% for the remaining five sessions of the study in Setting 4 and for four of the five remaining sessions in Setting 5. (See Figure 3.)

Training and Generalization: Subject 2

Training. Subject 2 reached criterion in Setting 1 in four training sessions on Targets 1, 2, and 3 (see Figures 5, 6, and 7). She took eight training sessions to reach criterion on Target 4 (see Figure 8).

Generalization. During baseline, there was erratic correct responding for Target 2 up to 50% during Sessions 2, 6, and 13 in Settings

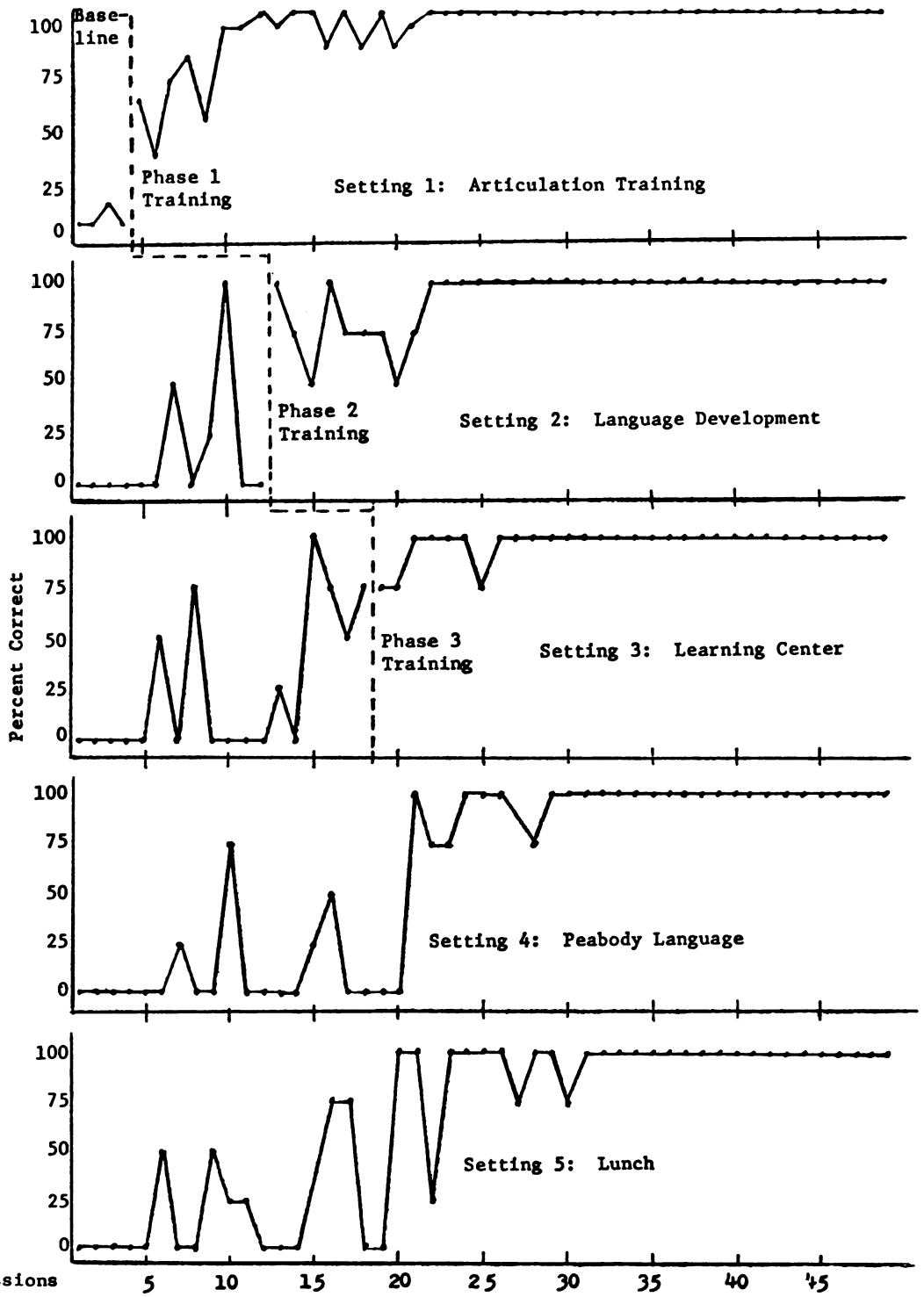

Fig. 1. Percentage correct responding on Target 1 for Subject 1. Broken line indicates when training was initiated. No training occurred in Settings 4 and 5.

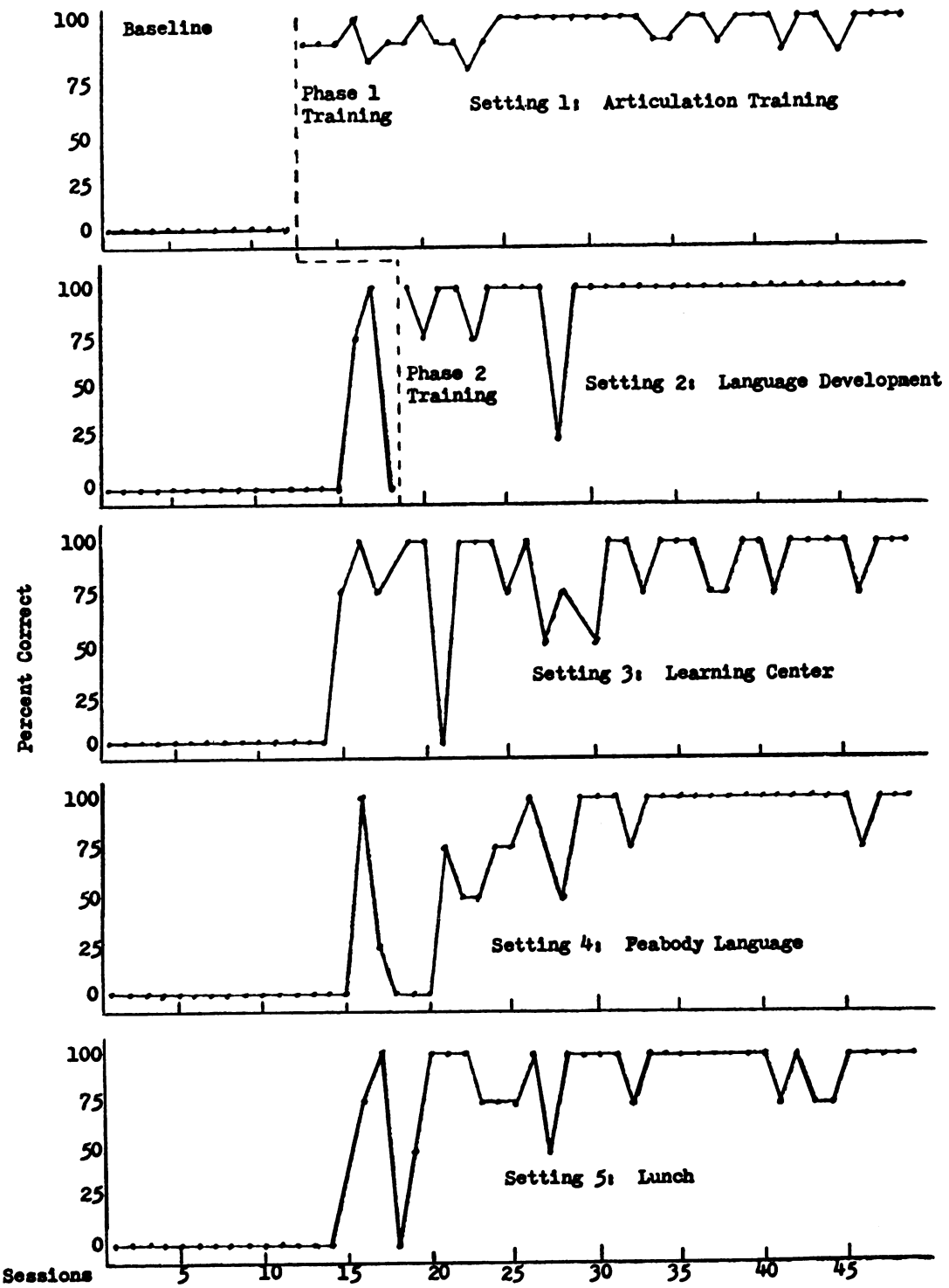

Fig. 2. Percentage correct responding on Target 2 for Subject 1. Broken line indicates when training was initiated. No training occurred in Settings 3, 4, and 5.

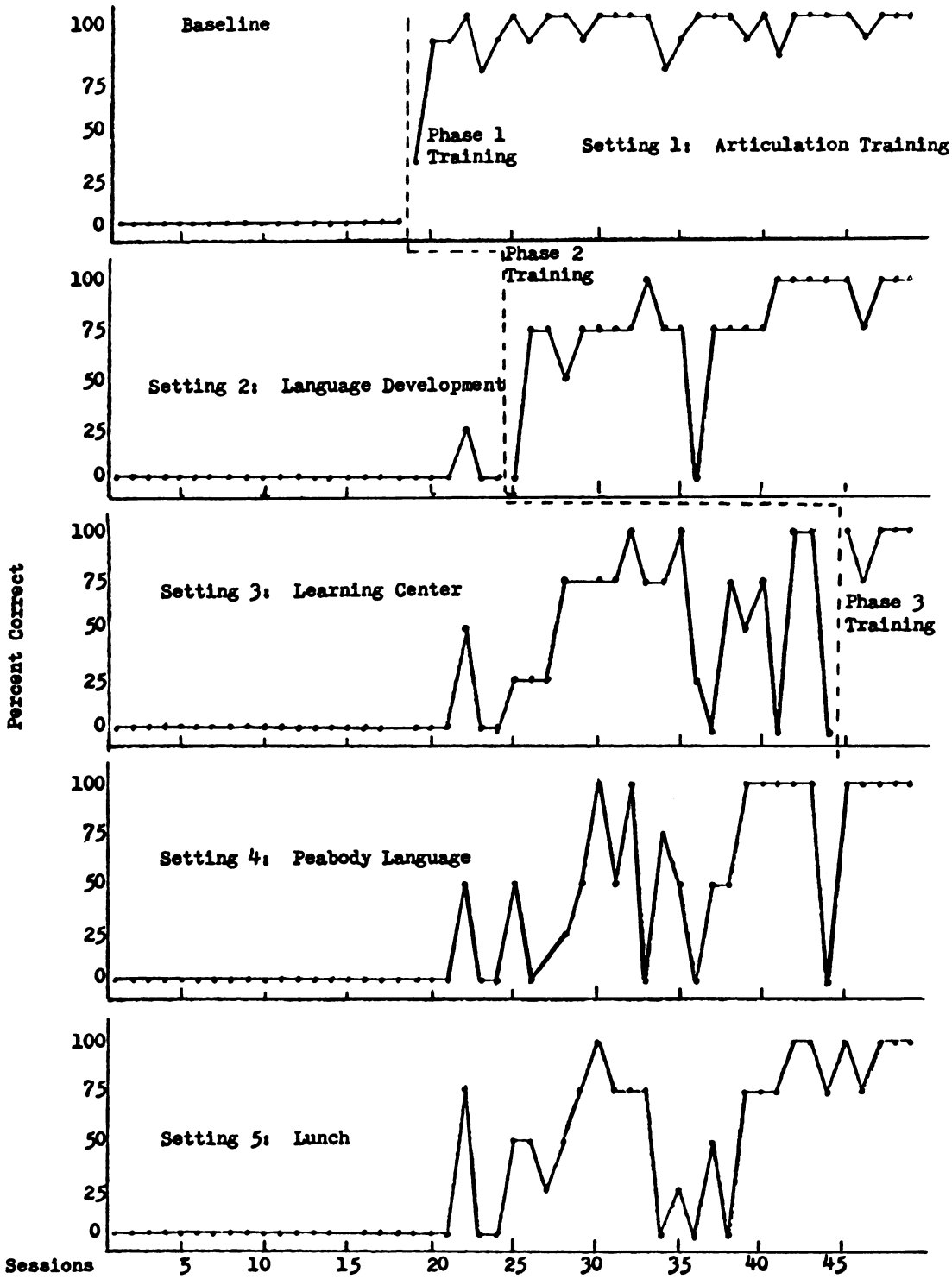

Fig. 3. Percentage correct responding on Target 3 for Subject 1. Broken line indicates when training was initiated. No training occurred in Settings 4 and 5.

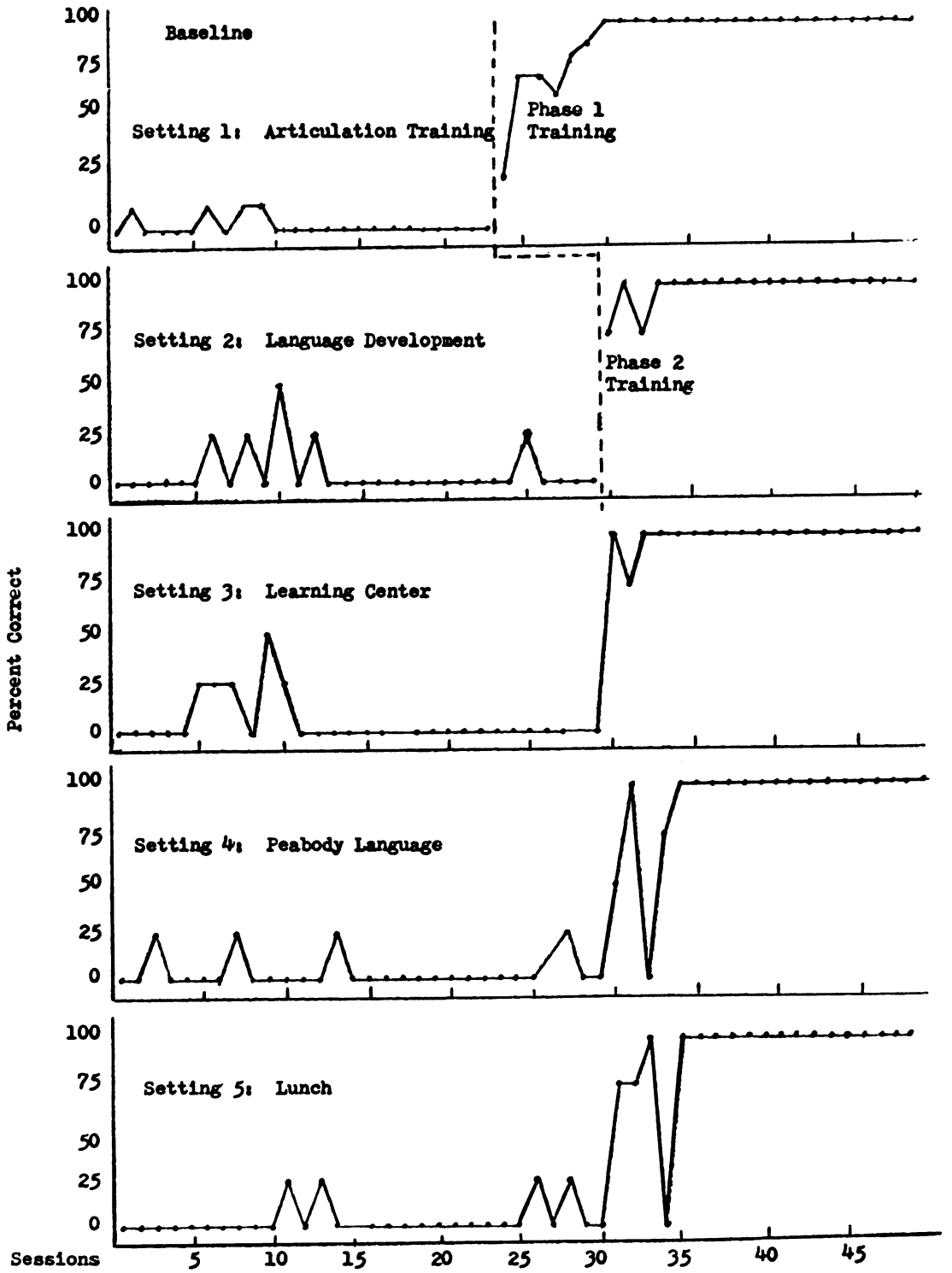

Fig. 4. Percentage correct responding on Target 4 for Subject 1. Broken line indicates when training was initiated. No training occurred in Settings 3, 4, and 5.

2, 3, and 4 (see Figure 6). Correct responding dropped to 0% in every setting during Session 14. Consequently, training was initiated in Setting 1, at which time correct responding for Target 2 reached criterion and stabilized in Settings 2, 3, 4, and 5 without necessitating training beyond the initial training setting.

During baseline, there was no correct responding in any setting for Targets 1, 3, and 4 (see Figures 5, 7, and 8). After training was initiated during Phase 1, there was one instance of correct responding on Target 1 in Setting 3. Correct responding for Target 1 remained at 0% throughout Phase 1 in Settings 2, 4, and 5 (see Figure 5). There was also one instance of correct responding on Target 4 in Setting 4. Correct responding for Target 4 remained at 0% throughout Phase 1 in Settings 2, 3, and 5. Consequently, training was initiated for Targets 1 and 4 during Phase 2 in Setting 2, after which correct responding for Target 1 reached criterion and stabilized in Settings 3, 4, and 5 at 100% for the remaining 38 sessions of the study, with two exceptions, 75% during Session 12 in Setting 4 and 75% during Session 33 in Setting 5. (See Figure 5.) Phase 2 was not initiated for Target 4 until Session 40 because of the multiple-baseline requirements of the study. Correct responding on Target 4 did not stabilize or reach criterion during the remaining nine sessions of the study in Settings 3 and 5. However, it did stabilize and reach criterion in Setting 4. (See Figure 8.)

After training was initiated during Phase 2, correct responding for Target 3 remained at 0% in Settings 3, 4, and 5 for three consecutive sessions. Consequently, training was initiated in Phase 3 in Setting 3, after which correct responding reached criterion and stabilized at 75% or better for the remaining 15 sessions of the study. (See Figure 7.)

DISCUSSION

The articulation training and generalization procedures in this study were effective in producing generalization of correct articulation in untrained settings and, in many instances, to persons not involved in the training. In almost every instance, generalization did not occur until training had occurred in at least one setting beyond the initial training setting. The single exception to this was Subject 2's second target word, "TV". The baseline data in Figure 6 suggest that this word was probably partially established beyond the imitation level before the onset of training. No other target word selected for this study resulted in similar unstable responding during baseline. Also, since Target 1 contained the /v/ in the final position and Target 2 contained the /v/ in the medial position, it is possible that this may have been an instance of similar task generalization. This did not occur for Targets 3 and 4 with Subject 2. However, these were the initial /r/ and the final vowelized /er/ which are, in fact, different phonemes. Similar task generalization in articulation is a variable that should receive further research, particularly related to whether or not it will also result in generalization to other settings and to other people. Raymore and McLean (1972) felt that training the phoneme in all three positions in words would preclude the need for situational generalization; Hartung (1970) stressed that speech must be trained in a number of settings in order for generalization to occur. Mowrer *et al.* (1968), Ryan (1971), and Costello and Bosler (1976) made situational training an integral part of their articulation programs by training parents. These three studies reported considerable success with articulation generalization.

The present results demonstrated that mentally retarded subjects can generalize correct articulation to untrained settings and to people not participating in the training. However, the investigation was limited to generalization of specific single-word responses that had been trained to criterion in the initial training setting. Further research is needed to determine whether a similar program would result in generalization to other untrained words containing the

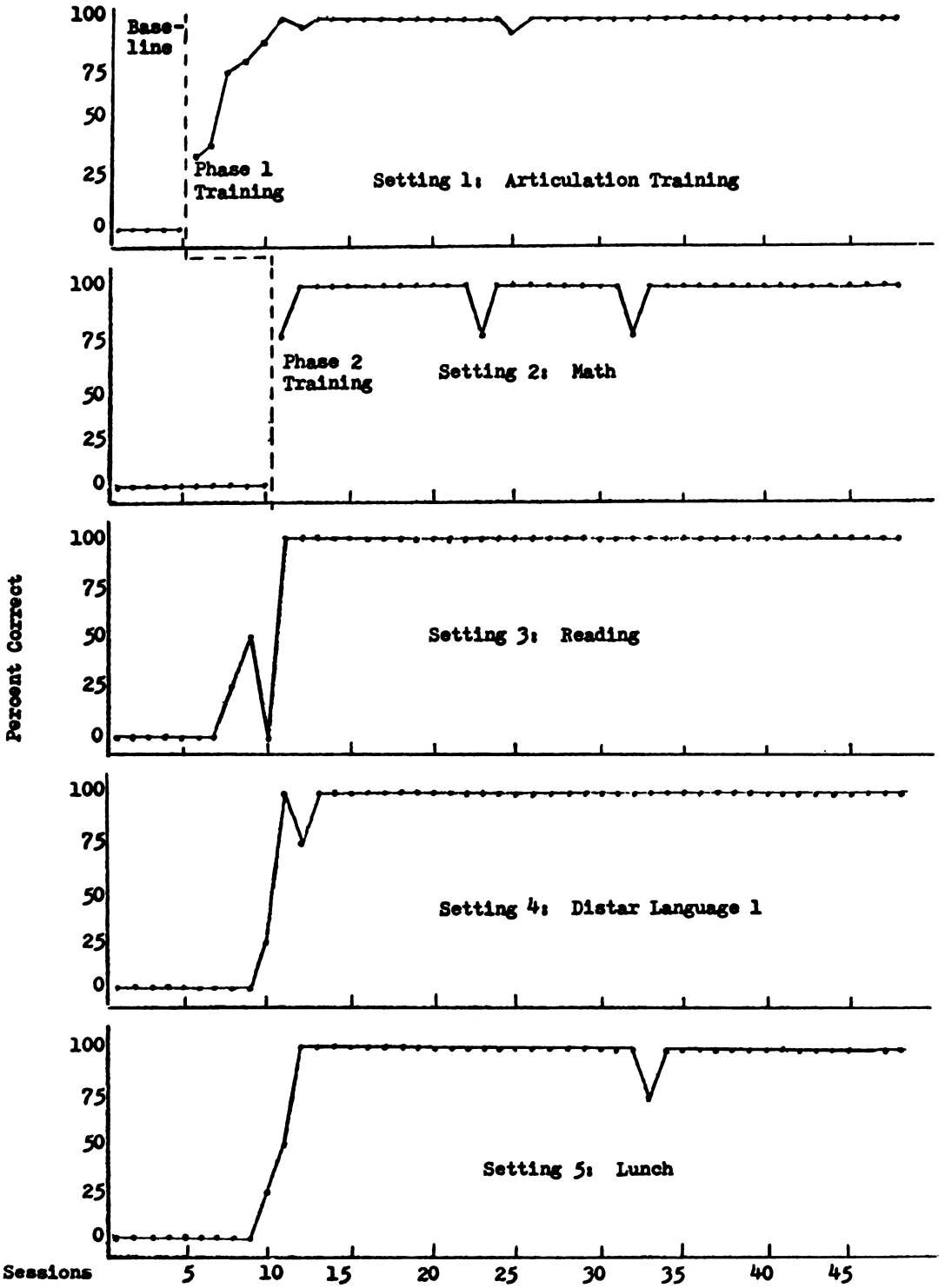

Fig. 5. Percentage correct responding on Target 1 for Subject 2. Broken line indicates when training was initiated. No training occurred in Settings 3, 4, and 5.

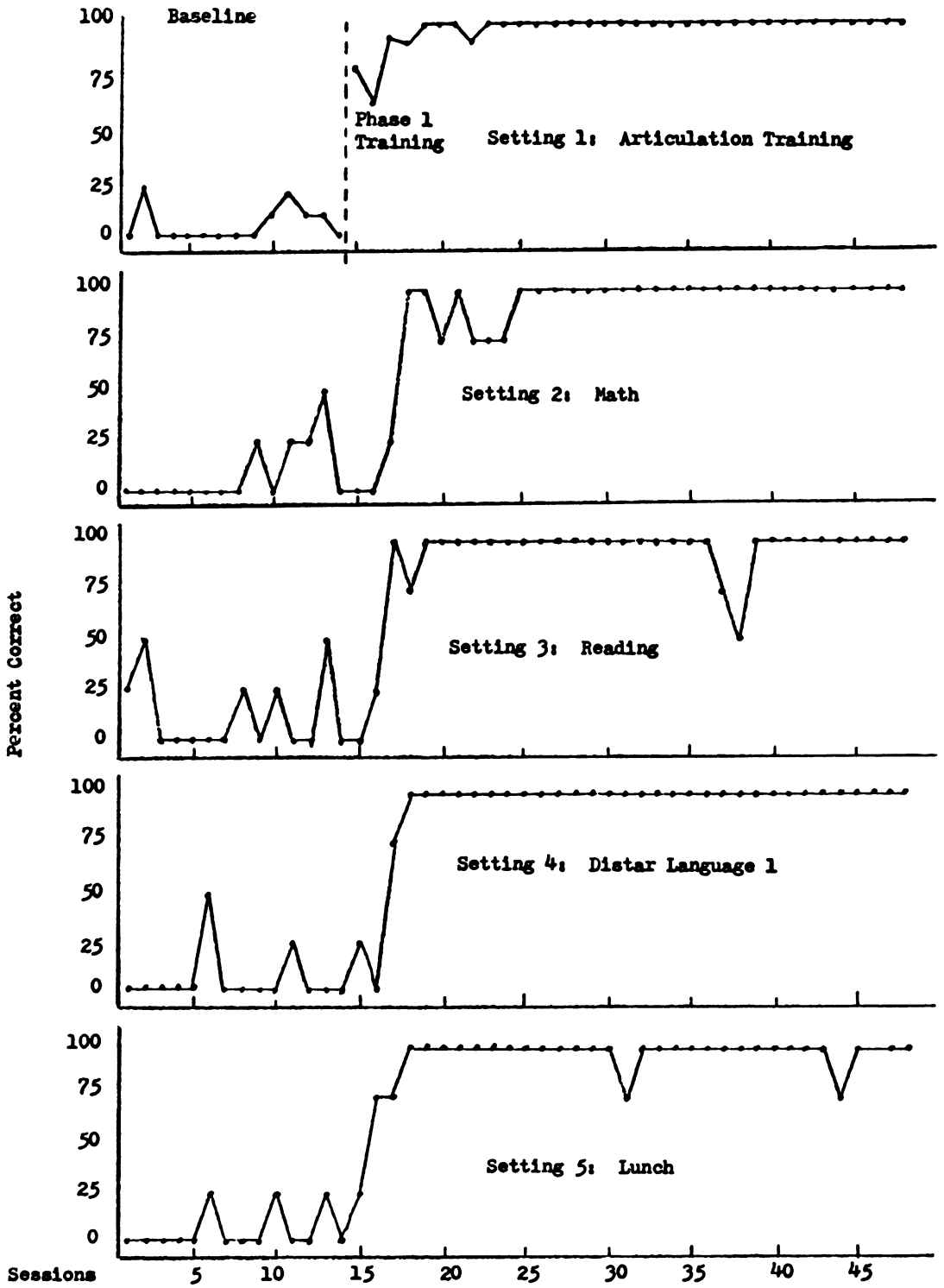

Fig. 6. Percentage correct responding on Target 2 for Subject 2. Broken line indicates when training was initiated. No training occurred in Settings 2, 3, 4, and 5.

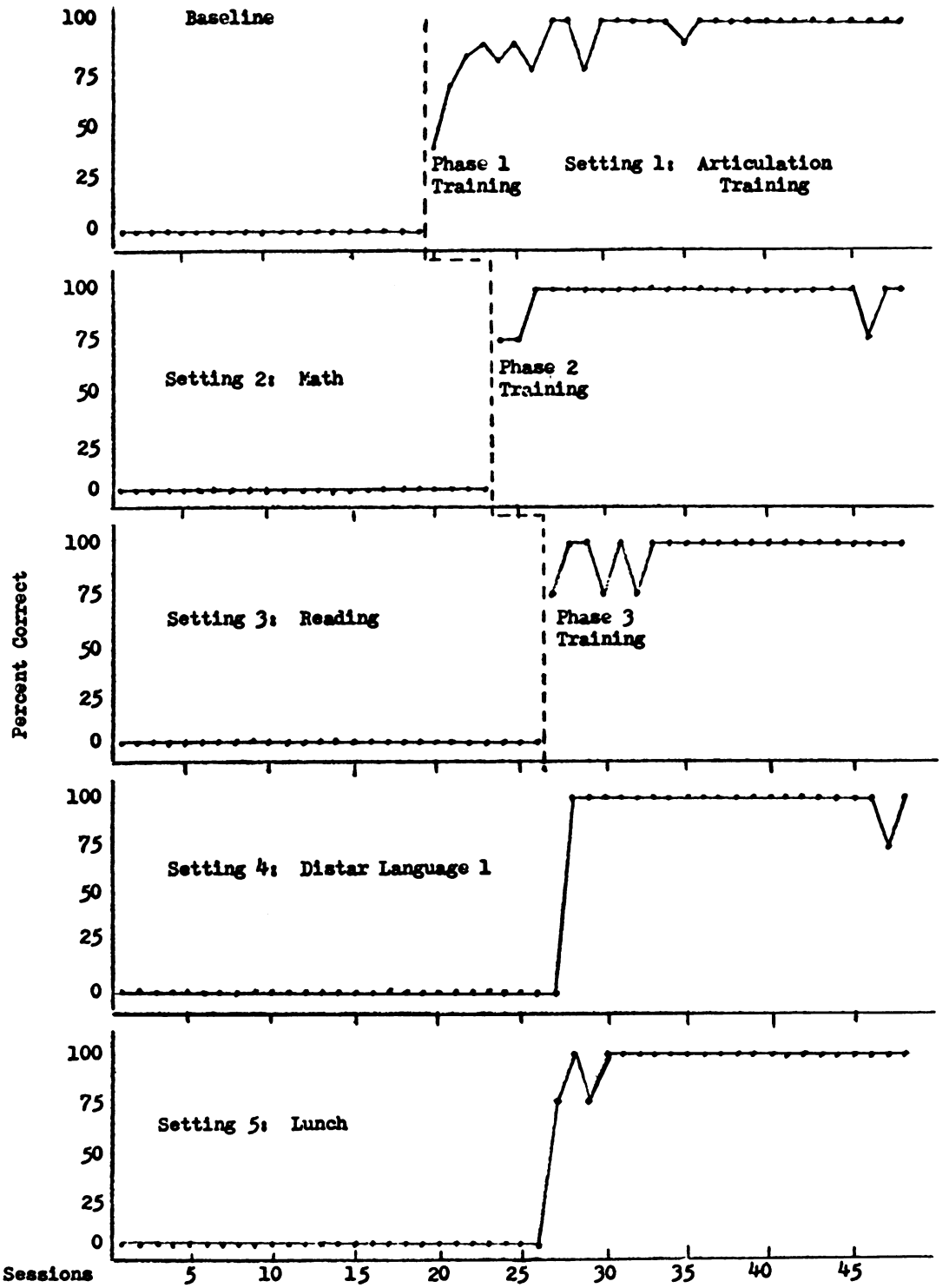

Fig. 7. Percentage correct responding on Target 3 for Subject 2. Broken line indicates when training was initiated. No training occurred in Settings 4 and 5.

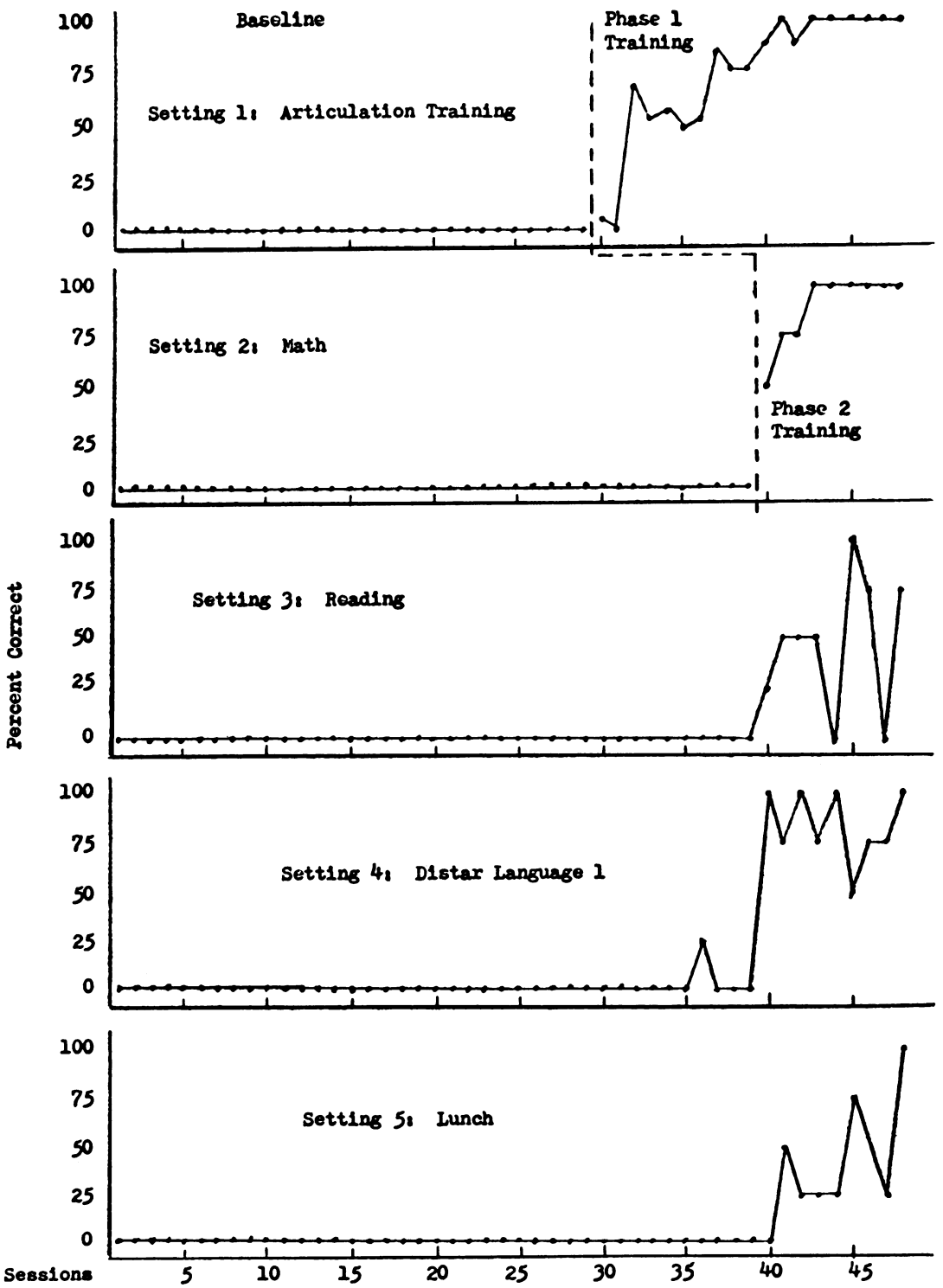

Fig. 8. Percentage correct responding on Target 4 for Subject 2. Broken line indicates when training was initiated. No training occurred in Settings 3, 4, and 5.

same phonemes and also whether it would result in correct articulation of the trained (and possibly untrained) words during connected discourse in addition to simple, single-word responses.

In addition, because the training was initiated in different settings and with different people doing the training simultaneously, it could not be determined whether one or both variables were necessary to accomplish the same generalization. There was also a simultaneous increase in the total amount of training occurring for each target word as training was initiated in a new setting. DeHaven and Garcia (Note 1) suggested that over-training accomplished generalization. Further investigation is necessary to isolate which variable or combination of variables are required to accomplish articulation generalization, as well as other forms of speech and language generalization.

Generalization cannot be taken for granted with any person, particularly the mentally retarded person. Consequently, it seems essential that speech correctionists include generalization measurements and training procedures in their articulation programs.

REFERENCE NOTE

1. DeHaven, E. and Garcia, E. *Continuation of training as a variable influencing the generalization of speech in a nonverbal retardate.* Paper presented at the meeting of the Rocky Mountain Psychological Association, Denver, May 1974.

REFERENCES

Baer, D. M. and Guess, D. Receptive training of adjectival inflections in mental retardates. *Journal of Applied Behavior Analysis,* 1971, **4,** 129-139.

Baer, D. M., Wolf, M. M., and Risley, T. R. Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis,* 1968, **1,** 91-97.

Costello, J. and Bosler, S. Generalization and articulation instruction. *Journal of Speech and Hearing Disorders,* 1976, **41,** 359-373.

Garcia, E. The training and generalization of a conversational speech form in nonverbal retardates. *Journal of Applied Behavior Analysis,* 1974, **7,** 137-149.

Guess, D. and Baer, D. M. Linguistic development in retarded children. In B. B. Lahey (Ed), *The modification of language behavior.* Springfield, Illinois: Charles C Thomas, 1973.

Guess, D., Sailor, W., Rutherford, G., and Baer, D. M. An experimental analysis of linguistic development: the productive use of the plural morpheme. *Journal of Applied Behavior Analysis,* 1968, **1,** 297-306.

Gray, B. B. and Ryan, B. *A language program for the nonlanguage child.* Champaign: Research Press, 1973.

Hartmann, D. P. Considerations in choosing an interobserver reliability estimate. *Journal of Applied Behavior Analysis,* 1977, **10,** 103-116.

Hartung, J. R. A review of procedures to increase verbal imitation skills and functional speech in autistic children. *Journal of Speech and Hearing Disorders,* 1970, **35,** 203-217.

Johnson, S. M. and Bolstad, O. D. Methodological issues in naturalistic observation: Some problems and solutions for field research. In L. A. Hammerlynck, L. C. Handy, and E. J. Mash (Eds), *Behavior change, methodology, concepts and practice.* Champaign: Research Press, 1973.

Mowrer, D. E. Transfer of training in articulation therapy. *Journal of Speech and Hearing Disorders,* 1971, **36,** 427-446.

Murdock, J. Y. and Hartmann, B. A language development program: *From imitative gestures to basic syntactic structures.* Salt Lake City: Word Making Productions, 1975.

Raymore, S. and McLean, J. E. Carryover of articulation therapy. In J. E. McLean, D. E. Yoder, and R. L. Schiefelbusch (Eds), *Language Intervention with the retarded.* Baltimore: University Park Press, 1972.

Risley, T. R. and Wolf, M. Establishing functional speech in echolalic children. *Behaviour Research and Therapy,* 1967, **5,** 73-88.

Sailor, W. Reinforcement and generalization of productive plural allomorphs in two retarded children. *Journal of Applied Behavior Analysis,* 1971, **4,** 305-310.

Schumaker, J. and Sherman, J. A. Training generative verb usage by imitation and reinforcement procedures. *Journal of Applied Behavior Analysis,* 1970, **3,** 273-387.

Sloan, H. N., Johnston, M. K., and Harris, F. R. Remedial procedures for teaching verbal behavior to speech deficient or defective young children. In H. N. Sloane, Jr. and B. D. MacAulay (Eds), *Operant procedures in remedial speech and language training,* 1968.

Snyder, L. K., Lovitt, T. C., and Smith, J. C. Language training for the severely retarded: Five years of behavior analysis research. *Exceptional Children,* 1975, **42,** 7-15.

VanRiper, C. *Speech correction: Principles and methods.* 4th ed.; Englewood Cliffs: Prentice Hall, Inc., 1963.

West, R., Ansberry, M., and Carr, A. *The rehabilitation of speech.* 3rd ed.; New York: Harper & Brothers, 1957.

Wing, D. M. and Heimgartner, L. M. Articulation carryover procedure implemented by parents. *Language, Speech and Hearing Services in Schools,* 1973, **4**, 157-173.

Received 4 November 1976.
(*Final acceptance 16 February 1977.*)