THE TRAINING AND GENERALIZATION OF A
CONVERSATIONAL SPEECH FORM IN
NONVERBAL RETARDATE¹

EUGENE GARCIA

UNIVERSITY OF UTAH

Differential reinforcement and imitation were used with two retarded children to train three sequential verbal responses associated with the display of a picture and questions related to that picture. Each response consisted of a three-word chain in sentence form; combined with verbal responses from the experimenter, this trained sequence formed a short conversational unit. Three experimenters measured the use of each sentence in settings different than the one in which training took place, and with pictures different than those used during training. Two types of generalization sessions were used: (1) General sessions, during which 10 pictures never used during training were displayed to the subject with reinforcement delivered on a noncontingent basis, and (2) Intermixed sessions, during which 10 pictures never used during training were displayed to the subject, but a picture having received training was also displayed, and correct responses to this picture were reinforced on a variable schedule. Both subjects learned the sentences being trained. However, little generalization was evident from this training when all experimenters conducted General probe sessions. Generalization occurred with one experimenter only after that experimenter conducted Intermixed probe sessions. Generalization to a third experimenter was then observed (*i.e.*, after the first two experimenters had conducted Intermixed probe sessions) without the use of Intermixed probe sessions by this third experimenter.

The procedures for training expressive speech have actually consisted of two major phases. The first phase is directed towards establishing or improving imitative verbal behavior. The child is taught to reproduce, verbally, anything presented to him (verbally) by the therapist (Baer, Peterson, and Sherman, 1967; Hewett, 1965; Lovaas, 1968; Risley and Wolf, 1967). But imitative speech in itself has limited value and may even reach a pathological level, as observed in echolalic children. Accordingly, the second major phase of speech development extends the control of vocal verbal behavior to objects, persons, and events in the environment.

The child is taught to label objects appropriately, to recognize relationships, to ask and answer questions, to put words together in sentence form, and to exhibit other significant, possibly subtle, but important attributes of effective verbal communication (Sloane, Johnston, and Harris, 1968; Lovaas, 1968).

The second phase of training discussed above, the establishment of functional speech, has received much experimental attention (Lovaas, 1968; Risley and Wolf, 1967; Sloane *et al.*, 1968). These experimental evaluations have indicated that speech and language skills can be trained using imitation, fading, and differential reinforcement. Although the establishment of speech is an important therapeutic goal, the use of speech in other contexts than those originally trained must be a concomitant goal. A characteristic of training must then be the generalized use of trained speech skills. This generalization can be defined both functionally and procedurally (Risley and Wolf, 1967). If de-

¹This research was sponsored by a Ford Foundation fellowship granted to the author while at the University of Kansas, and a state of Kansas research grant award to Kansas Neurological Institute. Appreciation is extended to Dr. Donald M. Baer for his direction and to the staff of Kansas Neurological Institute. Reprints may be obtained from Eugene Garcia, Department of Psychology, University of Utah, Salt Lake City, Utah 84112.

defined functionally, the term describes the control of specific verbalizations in situations or contexts different than those in which training took place. A procedural definition merely indicates by what means specific verbalizations are brought under the control of different situations or contexts. Defined either way, generalization of appropriate speech should be a goal of any speech development or remediation program.

Unfortunately, speech generalization has not received the experimental attention necessary to establish it as a reliably scheduled outcome of speech training. Reports have indicated that newly trained speech is used in other situations (Risley and Wolf, 1967; Sloane *et al.*, 1968; Wolf, Risley, and Mees, 1964), but experimental evaluation and reliable documentation of such reports have not been provided.

Research with withdrawn mental patients (Kale, Kaye, Whelan, and Hopkins, 1968) has documented the generalized use of trained (reinforced) verbal greetings. The use of these greetings with an individual who never participated in any type of training was not accomplished until a number of different individuals had participated in the specific training of these greetings. A recent study (Lyon, *unpublished*) has further reported and documented the generalization of a speech repertoire. Hospitalized schizophrenics were given reinforcement for answering a number of "everyday" questions ("What's your name?", "What are you doing?", *etc.*) in the presence of one experimenter and two other settings. Some transitory generalization was observed with (nontraining) experimenters in settings where reinforcement was never dispensed. But generalization increased substantially only after speech was reinforced in a few of the setting-experimenter combinations possible.

These studies provide an empirical base for the present experimental analysis of generalization. The present study attempted to identify procedures and relationships with respect to the generalization of trained vocal verbal behavior (speech). In addition, it demonstrated the ac-

quisition of a conversational unit of speech in subjects who previously did not exhibit this behavior.

METHOD

Subjects

One female, S1, age 12, and one male, S2, age 18, residents of the Kansas Neurological Institute, were chosen as subjects. Subjects showed no evidence of neurological damage, were classified as profoundly retarded with respect to their intellectual and social level, and were considered by psychological testers as "nonverbal". This nonverbal classification was not precisely descriptive, since each had a small repertoire of words (*e.g.*, "cookie", "candy", their name) and could imitate verbally.

Materials

Pictures of items such as camera, fireplace, teapot, *etc.* were cut from retail catalogues and used as stimuli for all experimental sessions (training and generalization trials)². These pictures ranged in size from 0.5 in. by 2 in. to 2 by 2 in. They were chosen to portray objects not likely to be labelled by the subjects. If a picture was correctly labelled during any generalization trial, it was replaced by another picture. One of a particular six pictures was always used during training trials; a pool of 90 other pictures was used during generalization trials.

Generalization Baselines: Response and Setting Definitions

The verbal behavior to be trained was a conversational sequence. For training purposes, it was divided into three separate segments. Table 1 lists this conversational unit and each of its three segments, together with the preceding experimenter response (cue) for each segment.

²A complete list of items used and their order of presentation will be sent upon request.

Table 1
Conversational Unit and Probe Trial Format

	<i>Experimenter Response</i>	<i>Subject Verbal Response</i>
Segment I	Display picture	R1: "What is that?"
Segment II	"This is a (<i>label</i>); what do you see?"	R2: "It's a (<i>label</i>)."
Segment III	"Do you want the (<i>label</i>)?"	R3: "Yes, I do."

Thus, the subject was taught three distinct sentence responses: (1) "What is that?", said in response to the experimenter's display of a picture; (2) "It's a (*label*)", said in response to the experimenter's question "This is a (*label*); what do you see?"; and (3) "Yes, I do", said in response to the experimenter's question, "Do you want the (*label*)?"

Three separate settings were established for each subject. These settings were defined as combinations of particular areas of the institution and specific experimenters present there:

The first experimenter, E1, a male, always saw each subject individually in a small (8 by 8 ft.) soundproof room.

The second experimenter, E2, a female, always saw each subject individually in their respective bedrooms in their living units.

The third experimenter, E3, a female, always saw each subject individually in a 12 by 12 ft. playroom located in another section of the institution. (During the course of the study it was necessary to replace E3 with another experimenter, a male, E'3, who continued to see each subject in this setting.)

Training

Four verbal responses were trained individually at different times during the study. The first response trained was the question, "What is that?" During a training trial, the experimenter displayed the picture and waited 10 sec for a response. If the subject incorrectly responded to a training trial (made an incorrect verbal response or did not respond within 10 sec), the experimenter modelled the correct response, then went on to the next trial. Thirty training trials were presented each session until

the learning criterion was reached.³ At least a 10-sec interval separated each trial. Training continued using the same picture until a three-part sequential learning criteria was reached: first, five consecutive correct responses while reinforcement was dispensed after each correct response (VR 1); then eight consecutively correct responses while reinforcement was dispensed, on the average, after every two correct responses (VR 2); and finally, 10 consecutively correct responses while reinforcement was dispensed, on the average, after every three correct responses (VR 3).

The second response trained was the answer to an experimenter's question: "This is a (*label*); what do you see?" The response trained was "It's a (*label*)".⁴ The training procedure and learning criterion were both similar to those of the first training, except that the experimenter presented the picture, then waited 5 sec after any subject response before providing the cueing stimulus, "This is a (*label*); what do you see?" (This delay should have prevented reinforcement of this response from chaining the presently trained response.)

The third trained response was a verbal answer to the experimenter's question, "Do you

³During some sessions, more than 30 training trials were presented if the subject was in the process of meeting one of the sequential criterion requirements. For example, if the subject had already responded correctly six times consecutively on VR 3 by the thirtieth trial, the training continued until he met criterion (10 consecutive responses), or responded incorrectly. Consequently, it was possible to extend any training session from one to nine trials.

⁴The required response was initially, "It's a (*label*)". Because of articulation problems in both subjects with the final *s* of *It's*, the final *s* requirement on this particular word was dropped.

want the (*label*)?" The response trained was, "Yes, I do". Training and criterion were the same as for the second training.

The fourth response trained was the chained combination of the previous three responses. This accomplished the goal of training a chain of verbal responses which, when combined with an experimenter's verbal response, could serve as an example of conversational speech.

Training sessions were always conducted by E1. The same six pictures were used individually during training sessions; that is, after the learning criterion was met with picture #1, picture #2 was used; after criterion was met with picture #2, picture #3 was used; . . . , after criterion was met with picture #6, picture #1 was used, *etc.*

Generalization

To examine means of establishing the generalized use of these newly trained sentences, each of the three experimenters conducted a separate generalization probe after the specified learning criterion had been met on each picture. Therefore, on the same day that a subject had met the training criterion, each experimenter saw that subject during a separate session to probe for generalization. The time of day of the probes varied, although the same sequence always was maintained: E3 saw the subject first, then E2, then E1. At least 30 min separated a training session from a generalization session. As mentioned previously, E3 was replaced by another experimenter (E'3) during the course of the study. Before this replacement, it was necessary to establish the functional equivalency of E3 and E'3. On seven generalization sessions (not all consecutive) before E3 was replaced, each subject was seen by E'3. Therefore, on these days, each subject was seen in a separate generalization session by four experimenters.

Two types of sessions were used to evaluate generalization.

(1) *General probe sessions.* Ten pictures, drawn from a pool of 90 and grouped in nine sets of 10 each, were displayed individually dur-

ing these sessions. The display of each picture was considered a probe trial. The particular 10-picture set used in a session was chosen from those sets that had not been used in any of the last six probe sessions (including probe sessions of all three experimenters) nor more often than any other set. During a probe trial, the experimenter displayed a picture in the following manner: the picture was shown to the subject for 10 sec; still displaying the picture, E said "This is a (*label*); what do you see?"; E then waited 10 sec (still displaying the picture) and said, "Do you want the (*label*)?", E then waited 10 sec more. Reinforcement during these sessions was provided on the average of once each minute, with the stipulation that reinforcement occur between probe trials and at least 10 sec after any verbal response emitted by the subject. Each experimenter used a timer to help determine these temporal factors.

(2) *Intermixed probe sessions.* Twenty training trials were conducted during these sessions. The same picture on which the subject had just met criterion was used and correct responses were reinforced on a VR 3 schedule. Intermixed among these training trials were 10 probe trials similar to those described above. These probe trials were intermixed among the training trials such that no two probe trials consecutively followed each other, nor did a probe trial follow a reinforced training trial.

SEQUENCE OF TRAINING AND GENERALIZATION PROBES

Pretraining

To assess the verbal responding of the subjects before training was initiated, each experimenter conducted an initial pretraining *General* probe session.

Training of R1: "What is that?"

1. Each experimenter conducted a *General* probe session (six probe sessions by each experimenter).

2. Because little generalization was observed, an attempt was made to produce generalization with one experimenter. Therefore, E1 conducted an Intermixed probe session; E2 and E3 continued to conduct General probe sessions.

3. In an attempt to produce generalization with yet another experimenter, E1 and E2 conducted Intermixed probe sessions; E3 continued to conduct General probe sessions.

Training of R2: "It's a (label)."

1. To ensure that it was not simply increased training that produced the previously observed generalization with R1, 12 pictures were trained to criterion with this response (the set of six pictures trained twice, thus doubling the training initially received by R1): a General probe session by each experimenter followed after the subject had met criterion on each picture.

2. Since no generalization occurred in the General probe sessions, an attempt was made to produce generalization with one experimenter. Therefore, E1 conducted Intermixed probe sessions; E2 and E3 conducted General probe sessions.

3. An attempt was made to produce further generalization (and replicate the procedures and effects shown with R1). E1 and E2 conducted

Intermixed probe sessions; E3 continued to conduct General probe sessions.

Training of R3: "Yes, I do."

For S1, the sequence of generalization probes was the same as during the training of R1. For S2, training continued only for six pictures, with General probe session following criterion attainment on each picture.

Training of R1 + R2 + R3: "What is that?"; "It's a (label)."; "Yes, I do."

For S1, the sequence of generalization probes was the same as that for R1 and R3. For S2, training continued for only six pictures, with General probe sessions by each experimenter following attainment on each picture.

(Table 2 and Table 3 list the above mentioned sequence of generalization probe sessions for each subject.)

Measurement and Reliability

During training trials, the experimenter scored each response as correct or incorrect. During probe trials, the experimenter wrote down all the subjects' verbal responses for each of the three segments of a probe trial. A score was tabulated for each of these three segments,

Table 2
S1: Sequence and Type of Generalization Probe Sessions for Each Experimenter

<i>Response Receiving Training</i>	<i>Probe Session No.</i>	<i>E1</i>	<i>Probe Types E2</i>	<i>E3</i>
Pretraining		General	General	General
R1: "What is that?"	1 through 6	General	General	General
R1: "What is that?"	7 through 10	Intermixed	General	General
R1: "What is that?"	11 through 18	Intermixed	Intermixed	General
R2: "It's a (label)."	19 through 30	General	General	General
R2: "It's a (label)."	31 through 36	Intermixed	General	General
R2: "It's a (label)."	37 through 42	Intermixed	Intermixed	General
R3: "Yes, I do."	43 through 48	General	General	General
R3: "Yes, I do."	49 through 54	Intermixed	General	General
R3: "Yes, I do."	55 through 60	Intermixed	Intermixed	General
R1 + R2 + R3	61 through 66	General	General	General
R1 + R2 + R3	67 through 72	Intermixed	General	General
R1 + R2 + R3	73 through 78	Intermixed	Intermixed	General

Table 3
S2: Sequence and Type of Generalization Probe Sessions for Each Experimenter

<i>Response Receiving Training</i>	<i>Probe Session No.</i>	<i>E1</i>	<i>Probe Types E2</i>	<i>E3</i>
Pretraining		General	General	General
R1: "What is that?"	1 through 6	General	General	General
R1: "What is that?"	7 through 10	Intermixed	General	General
R1: "What is that?"	11 through 18	Intermixed	Intermixed	General
R2: "It's a (label)."	19 through 30	General	General	General
R2: "It's a (label)."	31 through 36	Intermixed	General	General
R2: "It's a (label)."	37 through 42	Intermixed	Intermixed	General
R3: "Yes, I do."	43 through 48	General	General	General
R1 + R2 + R3	49 through 54	General	General	General

on the basis of five points possible for each segment: one point for each correct word, and one point for each correctly ordered two-word combination. For each segment of a probe trial, only the first three words of a verbal response were scored. A total probe session score was tabulated for each of the three segments across the 10 probe trials of that session. (Thus, a total of 50 points was possible each session for each segment.)

Because all training and probe sessions were recorded on tape, it was possible to have an independent observer score the recorded responses in the same way as the experimenter present in the session. Various individuals served as reliability observers. The independent reliability observer chose randomly at least one training session of each sequence of six (or 12) training pictures to score. In addition, they chose at random at least one generalization session by each experimenter in this same sequence. (Consequences provided during training trials were deleted from the tapes to decrease the probability of observer bias.) Percentages of agreement between the respective experimenter and the independent observers were computed for each session scored by an independent observer. For training sessions, reliability was computed as the percentage on which the two scorers agreed on correct or incorrect responses; the experimenter scored each response during the session while the observer scored from a tape

of that session. For probe sessions, reliability was defined as the percentage of agreement with respect to the number of points given on each of the three probe segments of each probe trial. Since each probe session consisted of 30 probe segments (10 each of the three probe segments), percentage of agreement was based on these 30 opportunities to agree. However, this method allowed the experimenter and the reliability observer to agree on the number of points for each probe trial segment without necessarily agreeing on the content of the segment (e.g., both *what* and *is* could be scored as one point). For this reason, percentage of agreement was also assessed for response content during generalization sessions, using the same basic agree-disagree technique described above.

RESULTS

Reliability

Interobserver agreement on the scoring of training trials within a single session ranged from 90% to 100%. This reliability estimate reflects the percentage of agreement between an independent observer and E1 on each of 13 training sessions for S1, and nine training sessions for S2 (the number of trials for each of these training sessions ranged between 23 and 38).

For each of the subjects, interobserver agreement on the scoring of probe trial segments

within a single generalization session was computed separately for each experimenter. Fourteen sessions by each of the three experimenters were independently scored for S1. Ten sessions by each of the three experimenters were independently scored for S2. For each of three sessions, reliability was assessed with respect to agreement on the number of points awarded on each probe trial segment and agreement on vocal content on each probe trial segment. Reliability ranged from 80 to 100%, with the means ranging between 89% and 97%.

Training

Figures 1 and 2 plot trials to criterion on successive training pictures for S1 and S2, respectively. It will be recalled that the learning criterion for each picture required that the least number of trials in which a subject could reach criterion was 23 (five consecutively correct on FR 1; followed by eight consecutively correct

on VR 2; followed by 10 consecutively correct on VR 3). In almost all cases, S1 met criterion near this minimal level. Only during the first three pictures in the training of R2 and the first two pictures in the training of R1 + R2 + R3 did trials to criterion range above 40. Similar training results are evident for S2: in almost all cases, criterion was met near the minimal level. The exception was during the training of R1 + R2 + R3 with the first three pictures. Trials to criterion for each of these pictures were more variable and greater (60 to 185), but dropped to near minimal level (30) during training of the next three pictures.

Generalization

Figures 3 and 4 present generalization results during the training of each verbal response. On any individual probe-trial segment, a total of five points was possible for a completely correct response. Each complete probe trial consisted

Fig. 1. Trials to criterion on successive training pictures for S1. Solid vertical lines indicate changes in the response undergoing training; dotted vertical lines indicate changes in generalization format during the training of a particular response.

Fig. 2. Trials to criterion on successive training pictures for S2. Solid vertical lines indicate changes in the response undergoing training; dotted vertical lines indicate changes in generalization format during the training of a particular response.

of three separate probe-trial segments, with 10 probe trials per session; thus, it was possible for a subject to score a total of 50 points on each of the three separate segment responses ("What is that?"; "It's a (*label*)."; "Yes, I do."). Figures 3 and 4 display the number of points of 50 possible on the three probe-trial segments for each experimenter over successive generalization probe sessions. The type of generalization session (General or Intermixed) is also designated for each experimenter.

Generalization results are presented for S1 in Figure 3. During the pretraining sessions, responding was low (0 to 10 points) on all three segment responses with all three experimenters. When training began on R1, and all

experimenters conducted General probe sessions, a slight increase (from zero to 10) resulted in R1 responses with E1; no change in R1 responding occurred with E2 and E3. When E1 conducted Intermixed sessions, a gradual increase in R1 responding occurred with E1 over these successive probe sessions (to a high of 36 points during the last sessions). No substantial change was evident in R1 responding with E2 and E3. When E2 conducted Intermixed sessions, an increase resulted in R1 with E2 (30 to 50 points per session) and E3 (10 to 30 points per session). R1 responding continued to increase (between 40 to 50 points per session) with E1. Training of R1 produced little generalization with any of the three experi-

menters while they conducted General probe sessions. When E1 conducted Intermixed sessions, generalization increased with E1; when E1 and E2 conducted Intermixed sessions, generalization increased with E2 as well as with E3. No changes were observed in R2 and R3 responding throughout these sessions.

Similar results were obtained during the training of R2 ["It's a (label)."]. A General

probe session was conducted by each experimenter after the subject had met the learning criterion on each of the first 12 pictures. Previously, R2 responding ranged between four and 10 points per session before and during the training of R1. During the first 12 pictures in the training of R2, a small increase was observed in R2 responses with E1. The high during any one session was 14 points (compared

GENERALIZATION

Fig. 3. Generalization results with all three experimenters for S₁. Long solid vertical lines indicate the training of a different response. Short solid vertical lines indicate changes in generalization format for a particular experimenter; short dotted vertical lines indicate changes in generalization format with another experimenter.

Fig. 4. Generalization results with all three experimenters for S_2 . Long solid vertical lines indicate the training of a different response. Short solid lines indicate changes in generalization format for that particular experimenter; short dotted vertical lines indicate changes in generalization format with another experimenter.

to a previous session high of 10 points). No change in R2 responses with E2 and E3 was evident. During the next six sessions, E1 conducted Intermixed sessions and R2 responding gradually increased over successive sessions, reaching a high of 32 points in one session (with a range of 13 to 32 points). After S1 had met criterion on each of the succeeding six pictures, E2 also conducted Intermixed sessions. An increase was observed in R2 responses with E2 (a high of 29 points in one session). A slight increase was also observed in R2 responses with E3. Before these six sessions, the range of R2 points per session with E3 was 0 to 10. When Intermixed sessions were conducted by both E1 and E2, the range of R2 responding increased to between nine and 16 points per session. During training of R2, R1 generalization decreased and remained low with all three experimenters, while no systematic changes were observed in R3 responding.

Results during the training of R3 replicate the findings observed during the training of R1 and R2: (1) generalization increased with E1 when E1 conducted Intermixed sessions; (2) generalization increased with E2 when E2 conducted Intermixed sessions; (3) generalization increased with E3 when E1 and E2 both conducted Intermixed sessions. No systematic change in R1 responding occurred during these probe sessions. Accuracy of R2 responding was higher during the first probe session with E1 and E2, but then dropped to near zero.

Although generalization was high for certain components during General sessions with E1 and E2, results during the simultaneous training of R1, R2, and R3 were similar to those reported during the individual training of each response.

For S1, the generalization was produced with four separate responses using the intermixing technique. With each response, intermixing was required with two experimenters before generalization to the third experimenter was accomplished. Although generalization to this third experimenter was observed, it was never at the

level observed with the two experimenters who conducted Intermixed sessions.

Figure 4 presents generalization results for S2. Again, the number of points of 50 possible for each probe segment response is plotted by successive generalization probe sessions. During pretraining sessions, there were no R1 or R3 responses to any experimenter, but 10 points of R2 responding resulted with each experimenter. During training of R1, a slight increase in R1 responding with each experimenter was seen during the first six sessions (a range of zero to 19 points per session with E1; one to 24 with E2; and six to 15 with E3). During these sessions, each experimenter conducted a General probe session. During the next six sessions, E1 conducted Intermixed sessions; E2 and E3 continued to conduct General sessions. An increase in R1 responses to E1 occurred; no systematic changes were evident in R1 responding with E2 and E3. For the next six sessions, E2 also conducted Intermixed sessions and an increase in R1 responding resulted with E2, as well as with E3. No change in R2 or R3 responding occurred with either experimenter during these six sessions.

During the training of R2 ["It's a (*label*)."], a comparison of R2 responding during training of R1 and during the first 12 probe sessions of R2 training indicated no systematic changes. During the next six sessions, E1 conducted Intermixed sessions and an increase in R2 responses occurred with E1. Intermixed sessions with E1 and E2 resulted in an increase for R2 responses with E2 and E3.

For S2, training of R1 and R2 resulted in: (1) little generalization with E1 until Intermixed sessions were conducted by E1; (2) little generalization with E2 until E2 conducted Intermixed sessions; and (3) little generalization with E3 until E1 and E2 both conducted Intermixed sessions. Generalization of R1 and R2 with E3 was never as great as that observed with E1 and E2. Generalization of R1 reoccurred with E2 and E3 during Intermixed sessions of E1 for R2.

Before the R3 training, R3 responding had remained at zero with all experimenters. During training of R3, an increase in R3 responding with all experimenters was observed. Similarly, increases in R1, R2, and R3 responding occurred when these three responses were trained simultaneously. Only General probe sessions were conducted by each experimenter during the training of these responses.

Equivalence of Experimenters

As mentioned previously, it became necessary during the course of the study to replace E3 with another experimenter (E'3). E'3 conducted a generalization session on seven of the same days in which E3 conducted a generalization session. A comparison of generalization with E3 and E'3 for both subjects indicated similar scores on those days in which the two conducted generalization sessions.

DISCUSSION

The acquisition of a trained verbal repertoire consisting of three separate sentence responses was observed in the behavior of both subjects of this study. Before this training, each subject's verbal repertoire consisted almost entirely of imitative speech. Consequently, control of speech was shifted from imitative to nonimitative cues (Lovaas, 1968; Risley and Wolf, 1967). The three separately trained sentence responses and the chaining of these same responses were trained using a combined procedure of imitation and differential reinforcement. By training these separate verbal responses simultaneously, each subject was in fact taught a short "conversational" speech form.

The generalized use of these trained responses was systematically observed to occur in contexts other than those directly involved during training. Generalization in this study refers to the use of a trained verbal response across either or both of the following dimensions: (1) pictures that were not directly used during training, (2) individuals (experimenters) who did not par-

ticipate in training. Since E1 conducted all training sessions, the generalization reported with this experimenter referred to correct responses to untrained pictures. Generalization reported for E2 reflected correct responses across the two dimensions during all General sessions. However, during Intermixed sessions, this experimenter was at least partially involved in a training procedure; thus, generalization was measured across only one dimension (untrained pictures). The generalization observed with the third experimenter occurred across both dimensions (untrained pictures and an individual who was never involved in training).

For each subject, a fixed sequence of training exposure occurred across experimenters. The exact effects of this standard order are confounded with the effects of the intermixing procedures. Although this constraint does not seem to negate the above conclusion, it is worthy of mention.

The intermixing procedure that resulted in generalized speech usage entailed concurrently (within the same sessions) reinforcing correct responding to some pictures while demonstrating other pictures to which responding was never reinforced. Similar intermixing procedures have in the past been employed in studies of traditional stimulus generalization (Terrace, 1966). Recently, this intermixing technique has been extended to the investigation of generalized imitation (Baer *et al.*, 1967; Baer and Sherman, 1964; Garcia, Baer, and Firestone, 1971) and generalized grammar (Guess, 1969; Schumaker and Sherman, 1970). These studies have intermixed nonreinforced (or nontrained) responses within reinforced responses in an effort to analyze generalization effects. Therefore, the present study draws on this previous foundation and adds to the systematic evaluation of this procedure as a technique that could be used to produce speech generalization.

Therapeutically, the subjects learned a verbal repertoire that was made to generalize to an individual who was never involved in the training. These same results were reported by Kale

et al. (1968) and Lyon (*unpublished*) and were discussed briefly earlier in this paper. The procedures by which each of these studies produced generalization share similarities with the present procedure. In each case, generalization of trained speech was not observed with non-training individuals until several individuals had conducted training on the same (or similar) verbal behavior. The present study differed somewhat in that an intermixing training procedure was utilized. Yet, a conclusion arrived at from all three studies might indicate that therapists involved in speech training should concomitantly program the generalized use of trained speech. Depending on the type of speech being trained, components of such a program might include training by more than one individual (in various contexts) and/or the use of an intermixing technique.

REFERENCES

- Baer, D. M., Peterson, R. F., and Sherman, J. A. The development of imitation by reinforcing behavioral similarity to a model. *Journal of Experimental Analysis of Behavior*, 1967, 10, 405-416.
- Baer, D. M. and Sherman, J. A. Reinforcement control generalized imitation in young children. *Journal of Experimental Child Psychology*, 1964, 1, 37-49.
- 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.
- Garcia, E., Baer, D. M., and Firestone, I. The development of generalized imitation within topographically determined boundaries. *Journal of Applied Behavior Analysis*, 1971, 4, 101-112.
- Guess, D. A functional analysis of receptive language and productive speech: acquisition of the plural morpheme. *Journal of Applied Behavior Analysis*, 1969, 2, 55-64.
- Hewett, F. M. Teaching speech to an autistic child through operant conditioning. *American Journal of Orthopsychiatry*, 1965, 35, 927-936.
- Kale, R. J., Kaye, J. H., Whelan, P. A., and Hopkins, B. L. The effects of reinforcement on the modification, maintenance, and generalization of social responses of mental patients. *Journal of Applied Behavior Analysis*, 1968, 1, 307-314.
- Lovaas, O. I. A program for the establishment of speech in psychotic children. In H. N. Sloane and B. D. MacAulay (Eds.), *Operant procedures in remedial speech and language training*. Boston: Houghton, Mifflin Co., 1968. Pp. 125-156.
- Lyon, V. L. *Conditioning and generalization of verbal response in hospitalized schizophrenics*. Unpublished doctoral dissertation, University of Kansas, 1971.
- Risley, T. R. and Wolf, M. M. Establishing functional speech in echolalic children. *Behaviour Research and Therapy*, 1967, 5, 73-88.
- Schumaker, J. and Sherman, J. A. Training generative verb usage by imitation and reinforcement procedures. *Journal of Applied Behavior Analysis*, 1970, 3, 273-287.
- Sherman, J. A. Use of reinforcement and imitation to reinstate verbal behavior in mute psychotics. *Journal of Abnormal Psychology*, 1965, 70, 155-164.
- Sloane, 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 and B. D. MacAulay (Eds.), *Operant procedures in remedial speech and language training*. Boston: Houghton Mifflin Co., 1968. Pp. 77-101.
- Terrace, H. S. Stimulus control. In W. K. Honig (Ed.), *Operant behavior: areas of research and application*. New York: Appleton-Century-Crofts, 1966. Pp. 271-344.
- Wolf, M. M., Risley, T. R., and Mees, H. I. Application of operant conditioning procedures to the behavior problems of an autistic child. *Behaviour Research and Therapy*, 1964, 2, 305-312.

Received 17 August 1972.

(Revision requested 6 December 1972.)

(Final acceptance 27 December 1973.)