*GENERALIZATION AND TRANSFER BETWEEN
COMPREHENSION AND PRODUCTION: A COMPARISON
OF RETARDED AND NONRETARDED PERSONS*

ANTHONY J. CUVO AND MARIA T. RIVA

SOUTHERN ILLINOIS UNIVERSITY AT CARBONDALE

The study compared the acquisition, generalization, transfer, and maintenance of language comprehension and production responses by persons at two IQ levels: mentally retarded ($N = 10$) and nonretarded ($N = 10$). The two levels of the IQ Level factor were combined factorially with two levels of a Training Condition factor: Comprehension-Production and Production Only. Participants in the former groups were trained sequentially to (a) comprehend coin labels by pointing, and then (b) produce verbally the correct coin label. Participants in the Production Only groups were trained on the latter response only. A three-factor mixed design with one repeated measure plus a multiple baseline across coin responses was employed. Results indicated that both mentally retarded and nonretarded subjects attained a high level of acquisition and maintained their performance on 1- and 4-week follow-up tests. No difference occurred between mentally retarded and nonretarded participants in magnitude of acquisition, but the mentally retarded groups took approximately three times as many trials to complete training. Data also suggested, contrary to past research, that generalization from comprehension to production was bidirectional, with no difference in magnitude between mentally retarded and nonretarded subjects. Transfer from comprehension to production occurred in both nonretarded and retarded subjects; comprehension training facilitated a savings of trials in production training. These results show that language differences between retarded and nonretarded persons are quantitative rather than qualitative as some past research may have suggested.

DESCRIPTORS: retardates, language, generalization, transfer

Language acquisition includes both comprehension and production components. Comprehension has been defined as the process by which the listener grasps the meaning of a word, phrase, sentence, or longer discourse (Carroll, 1972). Production is a linguistic response under stimulus control of some situation or environmental context (Seigel & Spradlin, 1978). A debate begins when experimenters and theorists attempt to determine which appears first in language acquisition: comprehension or produc-

tion. Traditionally, it has been suggested that auditory comprehension of the stimulus labels precedes verbal production. Early theory (Myklebust, 1957) and research (Fraser, Bellugi, & Brown, 1963; McCarthy, 1954) on language development confirmed this notion. Recently, however, conceptualizations of language have led to a skeptical view of the traditional theory. From a behavioral framework, there is no reason to assume that comprehension is related in a specific way to production. Comprehension involves an implicit or explicit response to a linguistic utterance, whereas production involves a linguistic response controlled by a specific stimulus or environmental situation. Because different responses and stimuli are present in the two performances, they could develop independently of each other. Hence, it is not inevitable that com-

The authors thank the children and staff of the Puka Preschool, Carbondale, Illinois, and A. L. Bowen Developmental Center, Harrisburg, Illinois, for their cooperation throughout this study. We also thank Robert Polvinale for his assistance. Requests for reprints should be sent to Anthony J. Cuvo, Behavior Modification Program, Rehabilitation Institute, Southern Illinois University, Carbondale, Illinois 62901.

prehension must precede production (Seigel & Spradlin, 1978). Some authors have assumed that both language comprehension and production are a reflection of underlying cognitive or linguistic competencies (Seigel & Spradlin, 1978). Advocates of this theory also do not believe that comprehension must necessarily precede production.

Guess (1969) taught the receptive use of the plural morpheme to two severely retarded individuals and concluded that language reception and production may be independent of each other. A follow-up study (Guess & Baer, 1973) supported the previous results when generalization again did not occur when mentally retarded persons were trained in one modality. Chapman and Miller (1975), on the other hand, found that some children used an appropriate form productively before they did so on a comprehension task. Several other studies have also concluded that production may precede comprehension (Keeney & Wolfe, 1972; Strohner & Nelson, 1974).

The issue of whether: (a) comprehension precedes production, (b) production precedes comprehension, or (c) the two are independent of each other has not been completely resolved. Some researchers, instead, have concentrated on the issue of whether instruction in one modality might facilitate acquisition in the other modality. Guess and Baer (1973) classified this type of learning as cross-modal generalization.

Productive articulation of nonretarded children has been shown to be improved substantially by receptive discrimination training in the case of nonsense words (Mann & Baer, 1971) and individual sounds (Winitz & Preisler, 1965). Ruder and his colleagues (Ruder, Hermann, & Schiefelbusch, 1977; Ruder, Smith, & Hermann, 1974) taught Spanish vocabulary items in a comprehension task to nonretarded preschool and school-age children. The results showed that initial comprehension training resulted in some verbal production. Other recent studies with nonretarded children have shown the same trend of facilitation from comprehen-

sion to production (Shipley, Smith, & Gleitman, 1969; Vasta & Teitelbaum, 1976; Whitehurst, 1977). Three other studies also show a facilitative effect from comprehension to production in nonretarded adults (Asher, 1972; Kushner & Winitz, 1977; Pimsleur, 1963). This body of literature suggests that prior training of comprehension facilitates acquisition of production by nonretarded children and adults.

In contrast, studies with mentally retarded individuals seem to suggest the converse order of facilitation: verbal (i.e., productive) pretraining of relevant dimensions enhanced motor (i.e., receptive) performance on match-to-sample concept formation (Hamilton, 1966) and geometric form discrimination tasks (Dickerson, Girardeau, & Spradlin, 1964). Miller, Cuvo, and Borakove (1977) examined generalization and transfer between comprehension and production in mentally retarded subjects. They found that teaching reception did not affect expression, but teaching expression resulted in almost total generalization to reception. The results also indicated that there were no savings in trials to criterion on expression acquisition after having been trained on reception. Smeets and Striefel (1976) also found that, with little exception, manual sign discrimination had little effect on the expressive usage of these signs by a retarded deaf girl. These studies using mentally retarded subjects, in contrast to those using nonretarded persons cited above, show facilitation from language production to comprehension.

In light of this research it is reasonable to ask whether there is actually a difference in the direction of generalization and magnitude of transfer between comprehension and production in the retarded and nonretarded populations. Is language acquisition by retarded individuals qualitatively different or merely slower than the nonretarded population? A number of authors have stated that skills required to comprehend sentences appear to develop in mentally retarded children in a similar sequence to that of nonretarded children, but at a much slower rate (Graham & Graham, 1971; Lackner, 1968;

Lenneberg, Nichols, & Rosenberg, 1964). Byrne (1959) reported that all children regardless of intellectual level followed the same sequence of normal speech development.

Consequently, the present research was designed to answer questions that have not yet been addressed and examine problems that existing research has not completely resolved. No single study has compared nonretarded and retarded subjects on acquisition, generalization, and transfer between comprehension and production using the same procedures for both groups. It is difficult to determine whether the differences in subjects, stimuli, and responses in the existing literature have contributed to the contradictory results regarding facilitation between comprehension and production. Also, these conclusions have been drawn from studies employing only one to four subjects (Guess, 1969; Guess & Baer, 1973; Mann & Baer, 1971; Smeets & Striefel, 1976).

The purposes of this experiment were to determine whether: (a) comprehension training would generalize and transfer to production, (b) production training would generalize to comprehension, (c) it would be more efficient (i.e., require fewer trials) to teach verbal production of stimulus labels directly or to teach auditory comprehension prior to production, and (d) there is a difference in the direction of generalization, magnitude of acquisition, transfer, and maintenance between mentally retarded and nonretarded individuals.

METHOD

Subjects

A total of 20 children participated in the present experiment; 10 were mentally retarded and 10 were of average to above average IQ. The nonretarded participants' mean mental age (MA) was 5 yr, 1 mo (SD = 1.18), their mean chronological age (CA) was 4 yr, 4 mo (SD = .68), and their mean IQ was 108 (SD = 9.06). There were 4 females and 6 males.

The 10 mentally retarded participants' mean

MA was 5 yr, 2 mo (SD = 1.18), their mean CA was 16 yr, 2 mo (SD = 2.92), and their mean IQ was 44 (SD = 12.38). In this group there was 5 males and 5 females. The Peabody Picture Vocabulary Test Form B, was used to obtain MA and IQ scores for all participants.

A large pool of volunteers was screened, and the 20 children were selected on their demonstrated skill at imitating the names of the five American coins in response to the experimenter's model, as well as their lack of skill at pointing to coins when named by the experimenter and producing coin labels in response to the experimenter pointing to and asking for their names.

The retarded and nonretarded children each were divided into two experimental groups. Within retarded and nonretarded groups, individual members were matched on their CA [retarded, $t(4) < 1.00$; nonretarded, $t(4) < 1.00$] and IQ scores [retarded, $t(4) < 1.00$; nonretarded, $t(4) < 1.00$]. One member of each pair was randomly assigned to each of the two experimental groups. Among the four retarded and nonretarded groups, children were matched on pretest scores [comprehension, $F(3, 16) = 1.21$; production, $F(3, 16) < 1.00$] as well as MA, $F(3, 16) < 1.00$. There were no significant differences on any of the matching variables.

Materials and Experimental Sessions

Training materials included 180 coins—36 each of five American coins (i.e., nickel, dime, quarter, half-dollar, and silver dollar). Five coins, one of each value, were glued approximately 5 cm apart in random order on 10 × 30.5 cm masonite strips. All coins were placed on the boards randomly with respect to head or tail orientation. The figures on the coins appeared equally often in four different degrees of rotation (i.e., 0°, 90°, 180°, 270°) with respect to the base of the figure. There were 36 strips of the 5 coins—10 for comprehension and production tests, 1 for modeling trials, and 25 for training and rehearsal trials. These same stimuli were used for both comprehension and production

training; however, all boards were rotated 180° from their comprehension training orientation for production training. During training, a 25-space board was used to provide feedback on correct responding.

Individual training sessions were conducted for approximately 20 min, five times per week in an office provided by the child's school or institution. The child and trainer, alone in the room, sat at a table with the instructor seated on the child's right side.

Experimental Design

The experiment combined both group and same subject design methodology to investigate several experimental questions. A three-factor mixed design with one repeated measure was employed. Between subjects factors were IQ Level (Mentally Retarded vs. Nonretarded), and Training Condition (Comprehension-Production vs. Production Only). Test trials was a within subjects factor; however, it was an incomplete factorial since Posttest I was not administered to the Production Only groups for either dependent variable. A group of retarded and a matched group of nonretarded children served in the Comprehension-Production condition in which auditory comprehension was taught first, immediately followed by production training. The other two matched retarded and nonretarded groups were placed in the Production Only condition, and trained on only the production procedure.

The experimental design also included a multiple baseline across coin responses strategy within each of the two between subjects conditions. Children were given the coin label comprehension and production measures initially as pretests. During the comprehension phase of Comprehension-Production training, children received a comprehension test (assessing performance on all five coins) after meeting the criterion for each of the five coins. This test permitted examination of whether or not acquisition of the comprehension responses for coin labels were independent of each other. The

Comprehension-Production group received both tests after completing comprehension training to assess acquisition of comprehension (i.e., post-test) and generalization from comprehension to production.

In the Production Only groups and the Production phase of Comprehension-Production training, children received a production test after meeting criterion for each of the five coins for purposes of the multiple baseline. When children completed production training, the coin label comprehension and production tests were readministered as posttests. For the Production Only condition, these tests measured acquisition of production, and generalization to the untrained comprehension responses. The two tests were readministered 1 and 4 weeks later as maintenance measures.

Principal Dependent Measures

Coin label comprehension test. This test required children to comprehend the five American coin labels by pointing to the target coin when the experimenter stated its name. The experimenter presented the 10 coin sets successively. For each set, all five coins were presented simultaneously in random order. The experimenter stated, "Point to the (coin name)." All responses were recorded immediately and no feedback regarding the correctness of responses was given.

A correct response required children to point to the target coin within 10 sec following the verbalization of that coin name by the experimenter. If 5 sec elapsed without a response, the experimenter repeated the instructions. Children were given 5 more sec to respond before the trial was considered incorrect. A correct response was recorded following a nonprompted correction of an error. An incorrect response was recorded if children pointed to a nontarget coin or if the 10 sec elapsed without a response.

After children responded to the five coins for the first trial, the second set of five coins was presented. The same procedures were used for trials 2-10. Because children were required to

respond to each of the 5 coins per trial, there were 10 responses to each coin per test. The percentage correct was calculated for each coin, and a total percentage correct was calculated for the Coin Label Comprehension Test.

Coin label production test. This test required children to produce the five American coin labels by verbally naming the coin in response to the experimenter pointing to it and asking, "What is the name of this coin?" The experimenter presented the 10 coin sets successively. For each set, all 5 coins were presented simultaneously. The experimenter pointed to each coin in the random arrangement, from left to right.

The procedures for administering and scoring the production test were identical to those of the comprehension test. The only acceptable responses were "nickel," "dime," "quarter," "half-dollar," and "silver dollar." Some children repeated the same coin name for each of the five coins, or alternately repeated two coin names across each coin set. When children responded correctly to a coin, using the above strategies, it was recorded as a correct response.

Procedure

Children in the Comprehension-Production group received comprehension and production training successively, while the Production Only group received only production training. For training in both modalities, coin labels were taught using a four-step sequential instruction program: (a) modeling trials, (b) training trials, (c) proficiency testing, and (d) rehearsal trials.

Modeling trials. The same masonite strip of five American coins was used for all modeling trials. The experimenter modeled the correct response first. In the comprehension condition she pointed to Coin #1 and said, "Here is a line of coins. This coin is the (coin name)." The experimenter kept her finger on the coin and said, "Now you do it. Put your finger on the (coin name)." For production training, children were told the name of the coin and asked, "What is the name of this coin?" Following a correct response, verbal praise was delivered.

If children did not respond, the experimenter said, "I will help you," and prompted the proper response either manually (comprehension training) or verbally (production training).

The modeling trial was terminated when the correct response was imitated without prompting. Only one coin name was trained per session.

Training trials. Immediately subsequent to the modeling trial for a coin, training trials were administered during the same session. Children were asked either to point to the target coin (comprehension training) or to state verbally the coin name (production training). The training trials required children to respond correctly to the target coin only; they were not asked to respond to the other four coins in that set. After responding correctly to the target coin, the next random set of five coins was presented. Following each correct response, verbal reinforcers were provided by the experimenter, and children moved the marker one space up the 25-space feedback board. When children completed 25 consecutively correct responses, they received a reinforcer (i.e., gum, candy, balloon, baseball card) that had been selected by them at the beginning of the training trials.

An incorrect response resulted in corrective feedback, and the experimenter moved the marker back to the starting position. When children responded incorrectly during comprehension training the experimenter pointed to the target coin and said, "No, this is the (coin name)." Then the children were asked either to point to (comprehension training) or to name (production training) the target coin. After correction, verbal praise was given, but the marker was not moved up the board.

Proficiency testing. After meeting the training performance criterion of 25 consecutively correct responses, either the comprehension or production test, depending on the training condition in effect, was administered. Proficiency criterion was .90 correct test performance for all coins previously trained. Failure to meet criterion resulted in recycling through the relevant model-

ing and training trials until criterion was attained.

Rehearsal trials. The session after children met the proficiency test criterion for Coin #2, the first two coins were presented concurrently for rehearsal to a criterion of five consecutively correct trials on each coin. During these trials, the coins to be rehearsed were trained randomly. Initially, the experimenter modeled the correct responses for all coins to be rehearsed, and children imitated according to the modeling trial procedure. The rehearsal trials were procedurally similar to the training trials, except that children were required to respond to more than one coin. Rehearsal trials were administered following proficiency testing of Coins #2, #3, and #4. The trials were cumulative in nature and included all coins whose labels had been taught to that point in time. The performance criterion was five consecutively correct responses for each coin rehearsed.

Reliability of Observation

A second observer, naive with respect to the experimental phase in effect, independently scored the test and training dependent measures throughout the experiment. Reliability checks occurred for 35 comprehension and 35 production tests across all phases of the experiment. Reliability checks were also administered on 21 training and 21 rehearsal trials. Reliability scores were computed using the following formula: $\# \text{ Agreements} / (\# \text{ Agreements} + \# \text{ Disagreements}) \times 100$. An agreement was recorded when the primary and secondary observer scored the same pointing or verbal response. A disagreement was scored when dissimilar responses were recorded by the observers. The percentage of agreement between the experimenter and observer was 100% for 106 of the 111 reliability checks; the lowest agreement was 94%.

RESULTS

Comprehension and Production:

Acquisition, Generalization, and Maintenance

Figure 1 shows the mean proportion of correct responses on the coin label comprehension and production pre-, post-, and follow-up tests for the four experimental groups. Acquisition, generalization, and maintenance were assessed by comparing results on specific tests initially by visual analysis of the figure and, subsequently, by statistical analysis.

Acquisition. A fundamental concern was whether instructional procedures produced acquisition of the comprehension and production responses after their respective training. For Comprehension-Production groups, acquisition of comprehension was assessed by comparing the Pretest and Posttest I data, shown in the slashed histograms of the upper two graphs of Figure 1. Acquisition of production was assessed for these groups by comparing their scores (solid histograms) immediately before production training (Posttest I) with those scores immediately after completing production training (Posttest II). These data are also found in the upper two graphs of Figure 1. For the Production Only groups, acquisition of production was measured by comparing the Pretest and Posttest II solid histograms in the lower two graphs of Figure 1. This visual analysis suggests a high level of comprehension as well as production acquisition for all four groups. After producing low pretest scores, all 20 children met the criterion of .90 correct for all five coins in comprehension and production modalities. These data suggest no mentally retarded-nonretarded differences in level of acquisition.

Generalization. It was of interest to determine whether generalization occurred between comprehension and production and whether the magnitude and direction of generalization dif-

Fig. 1. Mean proportion correct on pre-, post-, and follow-up comprehension and production tests for participants in the Mentally Retarded Comprehension-Production (MR, C-P), Nonretarded Comprehension-Product (NMR, C-P), Mentally Retarded Production Only (MR, PO), and Nonretarded Production Only (NMR, PO) conditions.

ferred between mentally retarded and nonretarded subjects. Comparing the production Pretest and Posttest I scores for Comprehension-Production groups in the upper two graphs of Figure 1, it was apparent that generalization from comprehension to production occurred in both mentally retarded and nonretarded subjects. The mentally retarded and nonretarded groups had production pretest scores of .11 and .07, respectively. However, following comprehension training, but prior to production training, both groups increased to .75 on production. Of the 10 children in these two groups, only four scored below the mean of .75. The range of Posttest I scores was .54 to .96.

In the mentally retarded and nonretarded Production Only groups, generalization occurred from production to comprehension even though comprehension was never directly trained. This generalization was assessed by comparing the Comprehension Pretest and Posttest II scores in the lower two graphs of Figure 1. The mentally retarded group scored a mean of .24 on the Pretest and a mean of .96 on Posttest II, and the nonretarded group scored .21 on the Pretest and .96 on Posttest II. Only one of the 10 participants scored below the mean of .90 (score = .82). Visual inspection of the figure shows substantial bidirectional generalization for both mentally retarded and nonretarded participants.

Maintenance. For all four groups, 1- and 4-week follow-up tests were measures of comprehension and production maintenance. Mean group scores on these tests did not drop below .90 for any of the four groups. Only 3 of the 20 participants fell below .90 on the 1-week follow-up test. For the 4-week follow-up test, 6 children fell below .90. Visual inspection of the figure shows high maintenance and no mentally retarded and nonretarded differences.

Statistical analysis. In order to supplement the visual analysis of Figure 1 presented above, the test data were subjected to statistical analysis. For each of the two dependent measures, comprehension and production, an ANOVA for a three-factor mixed design with one repeated

measure was performed. Since Posttest I was not administered to the Production Only groups for either dependent variable, the analysis was for an incomplete factorial. These ANOVAs provide information on acquisition, generalization, and maintenance of test performance. For each analysis, the Trials main effect was significant: Comprehension, $F(4, 55) = 924.67$; Production, $F(4, 55) = 503.72$, $p < .0001$ for each. Post hoc Tukey tests, $\alpha = .05$, showed that each pretest score was significantly lower than subsequently administered tests within each condition. This result confirms the visual analysis impression that significant acquisition, generalization, and maintenance obtained. Additionally, for the production dependent variable, there was significant acquisition in that modality after direct training (i.e., Posttest II significantly exceeded Posttest I). All other test trials did not differ from each other for either the retarded or nonretarded samples.

The only other significant effect for the two ANOVAs was a Training Condition main effect for the Comprehension dependent variable, $F(1, 16) = 16.98$, $p < .001$. Averaging across all other conditions, comprehension test performance was better for the Comprehension-Production Condition ($\bar{X} = 82.53$) compared to the Production Only Group ($\bar{X} = 76.20$). Most interestingly, there were no statistical interactions. This established that there was consistent performance between the levels of a factor with respect to the levels of the other two factors. Most importantly, there were no differences between the two IQ groups with respect to acquisition, generalization, and maintenance of either the comprehension or production responses in either of the two training conditions.

Transfer of Training

A primary purpose of the study was to determine whether there was transfer of training between comprehension and production. Transfer of training was defined as a savings of production training trials following comprehension training. Evidence of transfer would be shown

by children in the Comprehension-Production group requiring fewer production training trials than participants in the Production Only condition. To assess transfer, the number of trials required in each of the four conditions to achieve production mastery was compared. Table 1 shows the number of training and rehearsal trials required to complete production instruction for the mentally retarded and nonretarded groups. The nonretarded Comprehension-Production group required only a mean of 83.40 (i.e., 83.40 ± 0.00) trials compared to 314.40 (i.e., 286.40 ± 28.00) for the nonretarded Production Only group. All five children in the Comprehension-Production group completed the production portion in fewer trials than the five members of the Production Only group. The mentally retarded Comprehension-Production group also required fewer trials ($\bar{X} = 428.20$;

i.e., 415.00 ± 13.20) than the mentally retarded Production Only group ($\bar{X} = 927.80$; i.e., 855.60 ± 72.20). Two children in the Production Only group finished the program before the final child in the Comprehension-Production group completed training. A between groups comparison of the number of production trials shows that the nonretarded Comprehension-Production group took, on the average, only 26% of the production trials to reach criterion compared to their preschool peers in the Production Only group. In contrast, mentally retarded Comprehension-Production members attained the production criterion in 46% of the trials required by their retarded peers in the Production Only group.

Statistical analysis was undertaken to answer the question of whether or not significant positive transfer of training obtained and whether it occurred to the same degree for retarded and nonretarded persons. Because of extreme heterogeneity of variance within groups [Cochran Test, $C(4, 5) = .70$, $p < .05$], the individual subject trial data were subjected to logarithmic transformations. A 2×2 ANOVA showed a significant Training Condition main effect, $F(1, 16) = 6.46$, $p < .05$. Children who had acquired the comprehension responses initially required significantly fewer trials to learn the production responses, compared to participants who were instructed in the production modality only. The interaction between IQ Level and Training Condition was not significant indicating that acquisition of production in the Comprehension-Production groups was not any more rapid for nonretarded compared to retarded subjects.

These data show that there was a savings in verbal naming trials by previous mastery of pointing, but the magnitude of transfer did not differ statistically between IQ groups. On the whole, the nonretarded groups completed training before the mentally retarded groups. However, two mentally retarded children in each group finished before their nonretarded counterparts.

Table 1

Mean number of training and rehearsal trials per child to attain comprehension and production acquisition criteria for mentally retarded (MR) and nonretarded (NMR) children.

	NMR Comprehension- Production	NMR Production only
Mean Comprehension Rehearsal Trials	499.80	—
Mean Comprehension Rehearsal Trials	19.60	—
Mean Production Training Trials	83.40	286.40
Mean Production Rehearsal Trials	0.00	28.00
Mean Total Trials	602.80	314.40
	MR Comprehension- Production	MR Production only
Mean Comprehension Training Trials	1514.80	—
Mean Comprehension Rehearsal Trials	51.60	—
Mean Production Training Trials	415.00	855.60
Mean Production Rehearsal Trials	13.20	72.20
Mean Total Trials	1994.60	927.80

Efficiency of Training

Another question asked in this investigation concerned whether or not it would be more efficient in terms of total training trials to teach the terminal verbal naming behavior directly or to teach a possibly more rudimentary skill, comprehension by pointing, prior to the goal behavior. Table 1 shows mean comprehension rehearsal and training trials along with the comparable production information. The data show that the nonretarded Comprehension-Production group required a mean of 602.80 trials and the nonretarded Production Only group just 314.40 trials to complete their respective training programs. One of the five children in the Production Only group took more trials than two children in the Comprehension-Production group. The range for the Comprehension-Production group was 357 to 872 trials and the range for the Production Only group was 139 to 551 trials.

For the mentally retarded Comprehension-Production group, a mean of 1994.60 trials were required to complete training compared to only 927.80 in the Production Only condition. Two of the five children in the Comprehension-Production group finished before three members in the Production Only group. The range for the Comprehension-Production group was 456 to 4621 trials. The range for the Production Only group was 186 to 1743 trials.

A statistical analysis was undertaken to examine group differences in efficiency of training. Because of severe heterogeneity of variance in the trials dependent measure [Cochran Test, $C(4, 5) = 5.58, p < .01$], the individual subject trials data were transformed using logarithms. A 2×2 ANOVA, performed on these transformed data, showed significant main effects: IQ Level, $F(1, 16) = 5.77, p < .05$ and Training Condition, $F(1, 16) = 4.36, p = .05$. The interaction between these variables was not significant, $F(1, 16) < 1$. These results show that retarded persons took approximately three times as many training trials to achieve criterion than nonretarded persons. Also, teaching production

only was approximately twice as rapid, on the average, than teaching comprehension plus production. Therefore, it was more efficient to teach the terminal behavior directly than to teach a presumably more rudimentary response as a prerequisite.

Training time was also examined. The nonretarded Comprehension-Production group took a mean of 7 hr, 20 min (range of 3 hr, 40 min to 10 hr, 40 min) to finish the program, while the nonretarded Production Only group finished in just 3 hr, 40 min (range of 1 hr, 40 min to 5 hr, 40 min). The mentally retarded Comprehension-Production group also took much longer ($\bar{X}$ of 21 hr; range of 5 hr, 20 min to 46 hr, 20 min) than the mentally retarded Production Only group ($\bar{X}$ of 11 hr; range of 2 hr to 19 hr, 20 min).

Multiple Baseline Analysis

Figures 2 to 4 show data for the multiple baseline across responses strategy. Each data point represents the mean proportion of correct responses for the designated coin on the coin label comprehension or production tests. The data plotted for coins 1 to 5 are derived from the proficiency tests on which participants met the .90 correct criterion for all coins previously trained.

Figure 2 shows comprehension performance before, during, and following training for the nonretarded and mentally retarded Comprehension-Production groups. Inspection of the baselines shows progressively increasing slopes across the five pointing responses. Because the coins were trained successively, the probability of chance accuracy increased from .20 to 1.00 across the five coins. The final baseline data point for each coin corresponds quite closely to chance expectations for both retarded and nonretarded participants. Confounding factors (e.g., history) that may jeopardize the experiment's internal validity do not seem tenable. It is unlikely that extraneous factors would operate in a manner that would produce increments across successive

Fig. 2. Mean proportion of coin labels comprehended across test trials for nonretarded and retarded children in the Comprehension-Production condition.

Fig. 3. Mean proportion of coin labels produced across test trials for nonretarded and retarded children in the Comprehension-Production condition.

coins that were randomly ordered for each child, and those increments would correspond so closely to those expected by chance. Acquisition and maintenance are also illustrated in Figure 2.

The third figure shows verbal production scores for the two Comprehension-Production groups. These expressive language data were obtained after children were taught the same responses in the receptive modality. The baseline data show criterion or near criterion responding on coins 2 to 5 when tested after the first coin was trained. A reasonable inference is that comprehension training generalized to verbal production. Maintenance of the expressive responses is shown.

Figure 4 shows multiple baselines across production of coin responses obtained before, during, and following production training for the nonretarded and mentally retarded Production Only groups. The collection of these data was not preceded by comprehension training. In both the retarded and nonretarded samples, the baselines remained relatively stable prior to teaching each coin label; the productive responses increased substantially only after instruction, suggesting that they were relatively independent of each other.

It is informative to examine performance on coin label production when those responses are trained subsequent to learning comprehension (Figure 3) and without prior instruction in the other modality (Figure 4). A comparison of the baselines of the two figures suggests that previous comprehension training may have facilitated subsequent production performance. In the absence of prior comprehension training, the five productive responses were learned independently. When Figures 3 and 4 are compared, it appears that transfer and generalization from comprehension to production may account for the increased baseline slopes in the former figure. The multiple baselines for individual participants showed a close correspondence with the group data presented in Figure 3 and in Figure 4.

DISCUSSION

The principal question addressed by the present research was whether language behavior of mentally retarded individuals is qualitatively different, quantitatively different, or no different than that of the nonretarded population. Past research has suggested the possibility of an interaction between IQ and the direction of generalization between comprehension and production. Earlier studies have shown that for nonretarded subjects acquisition of comprehension facilitated performance on production (Asher, 1972; Kushner & Winitz, 1977; Mann & Baer, 1971; Pimsleur, 1963; Ruder et al., 1977; Ruder et al., 1974; Shipley et al., 1969; Vasta & Teitelbaum, 1976; Whitehurst, 1977; Winitz & Preisler, 1965). Research employing mentally retarded persons shows the converse effect: acquisition of production skills generalized to comprehension (Dickerson et al., 1964; Hamilton, 1966; Miller et al., 1977). Thus, much of the past research showed generalization to be unidirectional and in opposite directions for low and average IQ groups, suggesting a qualitative performance difference between them.

Since stimulus conditions and subject characteristics differed considerably across past experiments, it was impossible to draw meaningful conclusions regarding generalization differences between retarded and nonretarded persons from that work. The present experiment was designed to answer the experimental question using matched groups of retarded and nonretarded persons under standardized experimental procedures. Contrary to past research suggesting qualitatively different performance between mentally retarded and nonretarded individuals, the present research showed that generalization trends were similar for the two IQ groups. Generalization was bidirectional, and similar in magnitude for retarded and nonretarded persons. The low pretest production scores for both groups of participants rose substantially to a mean of .75, after comprehension training. Generalization from production to comprehension

Fig. 4. Mean proportion of coin labels produced across test trials for nonretarded and retarded children in the Production Only condition.

was even more impressive. After production training, the retarded and nonretarded groups both averaged .96 on the comprehension generalization test.

Bidirectional generalization does not seem an unreasonable outcome when one considers the procedures. In the comprehension task, the experimenter produced the verbal coin label as a discriminative stimulus for the participants' pointing response. For the production task, the experimenter pointed to the target coin as a discriminative stimulus for the participants' verbal response. Thus, the discriminative stimuli may have served as models for the responses not yet directly trained. Given the necessity of prompting comprehension by using the production response as a discriminative stimulus, and vice versa, it is interesting that generalization in past research was unidirectional.

Another purpose of the experiment was to investigate transfer between comprehension and production responses. An earlier study (Miller et al., 1977) showed no savings of trials in the acquisition of coin labels, after previous instruction on comprehension for retarded children. The present study, in contrast, showed transfer of coin names for both retarded and nonretarded samples. The production phases of training for Comprehension-Production conditions were completed in 26% and 46% fewer trials compared to the Production Only conditions for mentally retarded and nonretarded persons, respectively. These data show no qualitative differences between retarded and nonretarded persons.

The question of why bidirectional generalization and positive transfer of training occurred in this study but not in previous research is an interesting one to resolve. Some recent evidence may be pertinent. In a study similar to this one (Cuvo, Note 1), Training Condition (Comprehension-Production and Production Only) was combined factorially with Response Difficulty (one and two syllable words) using one set of Hebrew letters as stimuli for both difficulty levels and college students as participants. The principal finding was an interaction between

the two factors with respect to generalization and transfer. Generalization from comprehension to production occurred for the one syllable, but not the two syllable responses. In contrast, generalization from production to comprehension obtained for both one and two syllable responses. Thus, generalization was bidirectional when the responses were easier, but unidirectional when they were more difficult. Positive transfer of training occurred only for the one syllable but not the two syllable responses. When the procedures and results of the Miller et al. (1977) and the present study are considered in light of the Cuvo (Note 1) findings, the difference between the two experiments may be understandable. Learning the monetary values of coins (Miller et al., 1977) may be a more difficult task than acquiring their names; therefore, there was only unidirectional generalization and no transfer in the earlier study, but bidirectional generalization and positive transfer in the current experiment.

Another measure of learning, acquisition, was examined using two dependent variables, test performance and trials to criterion. The results showed that the comprehension and production posttest scores of mentally retarded and nonretarded subjects were equally high; however, the retarded subjects took approximately three times as many trials to produce that high level test performance. The learning rate of retarded and nonretarded children differed; however, both groups attained the same acquisition criterion.

Maintenance on the 1- and 4-week follow-ups was very high for both groups and did not differ between them. Therefore, there were neither qualitative nor quantitative differences between the retarded and nonretarded groups on this measure of performance.

These data lend support to past research that suggested that language processes of mentally retarded persons are similar to those of the nonretarded population, but their acquisition may occur at a slower rate (Byrne, 1959; Graham & Graham, 1971; Lackner, 1968; Lenneberg et al., 1964). No qualitative differences were uncov-

ered in the present study; however, acquisition was substantially slower for retarded children. Several reasons for the differences in rate of learning between retarded and nonretarded persons have been discussed previously. For example, Robinson and Robinson (1976), integrating research on retardation, stated that retarded individuals tend not to be alert to cues available that would help them solve problems. They also do not readily attend to relevant dimensions when solving discrimination-learning problems, and they fail to take into account outcomes of previous trials in formation of learning sets. It is possible that in the present study the difference in learning rate could have resulted from a failure by the mentally retarded groups to attend to relevant aspects of stimuli (i.e., size and figure on coins) as quickly as the nonretarded groups. Research also has shown a positive correlation between IQ and performance on visual discrimination tasks when the tasks are of intermediate difficulty (Zeaman & House, 1966). Therefore, when participants are matched on MA, any difference would be attributed to differences in IQ. Some of the nonretarded children in the present research had IQ scores 80 points higher than some retarded children, which could account for the differences in acquisition rate. These data seem to suggest that mentally retarded individuals learn in the same manner as the nonretarded population, but because of some deficiencies in discrimination and problem solving they need more trials to attain acquisition.

The applied significance of the present experiment lies in the utility of the generalization and transfer data for designing language instruction programs. The sequence of components within a language training program has been influenced by the earlier theory and research suggesting that comprehension precedes production in normal development; therefore, it is common practice to teach comprehension before production. The current research shows that it may be more efficient to teach production directly. It may be unnecessary to take the time to teach

comprehension first on the assumption that it will facilitate acquisition of production.

REFERENCE NOTE

1. Cuvo, A. J. *Effect of response difficulty on generalization and transfer between language comprehension and production.*

Manuscript submitted for publication, 1979.

REFERENCES

- Asher, J. Children's first language as a model for second language learning. *Modern Language Journal*, 1972, **56**, 133-138.
- Byrne, M. C. Speech and language development of athetoids and spastic children. *Journal of Speech and Hearing Disorders*, 1959, **24**, 231-240.
- Carroll, J. B. Defining language comprehension. In R. O. Friedle and J. B. Carroll (Eds.), *Language comprehension and the acquisition of knowledge*. Washington, D.C.: V. H. Winston, 1972.
- Chapman, R., & Miller, J. Word order in early two and three word utterances: Does production precede comprehension? *Journal of Speech and Hearing Research*, 1975, **18**, 355-371.
- Dickerson, D. J., Girardeau, F. L., & Spradlin, J. F. Verbal pretraining and discrimination learning by retardates. *American Journal of Mental Deficiency*, 1964, **68**, 476-484.
- Fraser, C., Bellugi, V., & Brown, R. Control of grammar in imitation, comprehension, and production. *Journal of Verbal Learning and Verbal Behavior*, 1963, **2**, 121-135.
- Graham, J. T., & Graham, L. W. Language behavior of the mentally retarded: Syntactic characteristics. *American Journal of Mental Deficiency*, 1971, **75**, 623-629.
- 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.
- Guess, D., & Baer, D. M. An analysis of individual differences in generalization between receptive and productive language in retarded children. *Journal of Applied Behavior Analysis*, 1973, **6**, 311-329.
- Hamilton, J. Learning of a generalized response class in mentally retarded individuals. *American Journal of Mental Deficiency*, 1966, **71**, 100-108.
- Keeney, T. J., & Wolfe, J. The acquisition of agreement in English. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 698-705.
- Kushner, D., & Winitz, H. Extended comprehension practice applied to an aphasic patient. *Journal of Speech and Hearing Disorders*, 1977, **42**, 296-306.

- Lackner, J. R. A developmental study of language behavior in retarded children. *Neuropsychologica*, 1968, 6, 301-320.
- Lenneberg, E. H., Nichols, I. A., & Rosenberg, E. F. Primitive stages of language development in mongolism. In D. M. Rioch & E. A. Weinstein (Eds.), *Disorders of communication* (Research publications of the Association for Research in Nervous and Mental Disease, Vol. 42). Baltimore: Williams & Wilkins, 1964.
- Mann, R., & Baer, D. M. The effects of receptive language training on articulation. *Journal of Applied Behavior Analysis*, 1971, 4, 291-298.
- McCarthy, D. Language development in children. In L. Carmichael (Ed.), *Manual of child psychology*. New York: John Wiley & Sons, 1954.
- Miller, M. A., Cuvo, A. J., & Borakove, L. Teaching naming of coin values—Comprehension before production versus production alone. *Journal of Applied Behavior Analysis*, 1977, 10, 735-736.
- Myklebust, H. R. *Auditory disorders in children*. New York: Grune & Stratton, 1957.
- Pimsleur, P. Discrimination training in the teaching of French pronunciation. *Modern Language Journal*, 1963, 47, 199-203.
- Robinson, N., & Robinson, H. *The mentally retarded child*. New York: McGraw-Hill, 1976.
- Ruder, K., Hermann, P., & Schiefelbusch, R. Effects of verbal imitation and comprehension training on verbal production. *Journal of Psycholinguistic Research*, 1977, 6, 59-71.
- Ruder, K., Smith, M., & Hermann, P. Effects of verbal imitation and comprehension on verbal production of lexical items. In L. V. McReynolds (Ed.), *Developing systematic procedures for training children language*. ASHA Monographs, 1974, 18, 15-29.
- Seigel, G., & Spradlin, J. Programming for language and communication therapy. In R. L. Schiefelbusch (Ed.), *Language intervention strategies*. Baltimore: University Park Press, 1978.
- Shipley, E., Smith, C., & Gleitman, L. A study in the acquisition of language: Free responses to commands. *Language*, 1969, 45, 322-342.
- Smeets, P. M., & Striefel, S. Acquisition and cross modal generalization of receptive and expressive signing skills in a retarded deaf girl. *Journal of Mental Deficiency Research*, 1976, 20, 251-259.
- Strohner, H., & Nelson, K. The young child's development of sentence comprehension, influence of event probability, nonverbal context, syntactic form, and strategies. *Child Development*, 1974, 45, 567-576.
- Vasta, R., & Teitelbaum, M. Reception training effects on the production of modeled language constructions. *Journal of Experimental Child Psychology*, 1976, 22, 67-72.
- Whitehurst, G. Comprehension, selective imitation, and the CIP hypothesis. *Journal of Experimental Child Psychology*, 1977, 23, 23-38.
- Winitz, H., & Preisler, L. Discrimination pretraining and sound learning. *Perceptual Motor Skills*, 1965, 20, 905-916.
- Zeaman, D., & House, B. The relation of IQ and learning. In R. M. Gagne (Ed.), *Learning and individual differences*. Columbus, Ohio: Charles E. Merrill Books, 1966.

Received March 19, 1979

Final acceptance August 31, 1979