Perception & Psychophysics
1984, 36 (2), 131-145

Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English

WINIFRED STRANGE
University of South Florida, Tampa, Florida

and

SIBYLLA DITTMANN
University of Minnesota, Minneapolis, Minnesota

Native Japanese speakers learning English have difficulty perceptually differentiating the liquid consonants /r/ and /l/, even after extensive conversational instruction. Using a same-different discrimination task with immediate feedback, eight adult female Japanese were given extensive training on a synthetic "rock"-"lock" stimulus series. Performance improved gradually for all subjects over the 14 to 18 training sessions. Comparisons of pretraining and posttraining categorical perception tests with the training stimuli indicated transfer of training to the more demanding identification and oddity discrimination tasks for seven of the eight subjects. Five of seven subjects also improved in identification and oddity discrimination of an acoustically dissimilar "rake"-"lake" synthetic series. However, transfer did not extend to natural speech words contrasting initial /r/ and /l/. It was concluded that modification of perception of some phonetic contrasts in adulthood is slow and effortful, but that improved laboratory training tasks may be useful in establishing categorical perception of these contrasts.

In the phonological system of a particular language, only a subset of all possible phonetic contrasts are utilized to differentiate lexical items. Different languages utilize different subsets of phonetic contrasts; thus, the patterns of speech production and perception for a speaker-hearer are constrained by the phonological system of his or her native language. It has long been known that, for adults, learning a new pattern of speech *production* in a foreign language is problematic. The persistence of a foreign accent includes, among other things, a failure to produce phonetic contrasts appropriately in the new language, contrasts that are not distinctive (i.e., not utilized to distinguish lexical items) in the learner's native language. Both anecdotal evidence and experimental investigations suggest that similar problems seem to persist in speech *perception* as well (Briere, 1966; Goto, 1971; Nemser, 1971).

Cross-language differences in the perception of a variety of contrasts among both consonants and vowels have been demonstrated using natural speech stimuli (i.e., minimal pair contrasts produced by native speakers) and several sorts of perceptual tasks (Gottfried, 1984; Marckwardt, 1944, 1946; Sapon & Carroll, 1958; Singh & Black, 1966; Trehub, 1976; Werker, Gilbert, Humphrey, & Tees, 1981). Studies of categorical perception have also demonstrated cross-language differences in the perception of acoustic dimensions underlying one or more phonetic contrasts. Differences in the perception of voicing and aspiration contrasts in initial stop consonants, distinguished by the (acoustically complex) voice onset time (VOT) dimension, have been found among native speakers of English, Thai, Spanish, French, and Polish (Abramson & Lisker, 1967; Caramazza, Yeni-Komshian, Zurif, & Carbone, 1973; Keating, Mikos, & Ganong, 1981; Williams, 1977). In general, contrasts that were distinctive in the listeners' native language were easily discriminated and labeling functions had sharp category boundaries, whereas "foreign" contrasts were discriminated and labeled poorly. Beddor and Strange (1982) showed that

---

This research was supported by a grant to James J. Jenkins and Winifred Strange, "Studies of Speech Perception" (MH-21153), and by grants to the Center for Research in Human Learning from NICHHD (HD-01136) and NSF (BNS-772-22075). Preparation of the manuscript was supported by MH-37924. Portions of this work were reported at the 7th Annual Minnesota Regional Conference on Language and Linguistics, Minneapolis, 1981, and at the 22nd annual meeting of the Psychonomic Society, November 1981. The authors wish to express their appreciation to Lenief Heimstead and Grant Miller for developing the computer hardware and software, to James J. Jenkins and Terry Gottfried for their comments on the manuscript, and to the Japanese women who served so conscientiously as subjects in this experiment. S. Dittmann is associated with the Center for Research in Human Learning, at the University of Minnesota. Requests for reprints should be sent to Winifred Strange, Department of Communicology, University of South Florida, Tampa, FL 33620.

Americans had difficulty differentiating synthetic oral and nasal vowels /ba/-/bã/, while Hindi speakers, for whom the contrast is distinctive, perceived the contrast categorically. (See, however, Stevens, Liberman, Studdert-Kennedy, & Öhman, 1969.) Cross-language studies of the American English (AE) /r/-/l/ contrast will be reviewed in detail below.

Despite this growing body of research demonstrating language-specific patterns in the speech perception of adults, there have been almost no studies that have explored the ontogeny of these differences in first-language learning (Oller & Eilers, 1983; Werker, 1982) and only a few studies have documented how perceptual abilities change during second-language learning in children and adolescents (Streeter & Landauer, 1976; Williams, 1979). Recent cross-sectional studies have examined differences in phonetic perception in adults learning a foreign language (Flege, in press; Gottfried, 1984; MacKain, Best, & Strange, 1981). Results suggest that although intensive conversational instruction (with native speakers) is correlated with improved perception of the foreign contrasts, perception is still not native-like, even for the most advanced students. Thus, for adults learning a foreign language, modification of phonetic perception appears to be slow and effortful, and it is characterized by considerable variability among individuals.

There have been attempts to modify phonetic perception in the speech laboratory, all concentrating on the VOT dimension underlying voicing and aspiration contrasts in initial stop consonants. Early studies that used absolute identification (Lisker, 1970; Strange, 1972) or variable-standard discrimination (Pisoni, 1971; Strange, 1972) with or without feedback met with relatively little success. Pisoni, Aslin, Perey, and Hennessy (1982) recently reported a study in which about half of the AE speakers tested were able to differentiate synthetic voiced vs. voiceless unaspirated initial stops (a non-English contrast) with little or no training. In a second experiment, they found no improvement in discrimination performance as a function of immediate feedback. In a third study, training in categorization of clear cases of each of the three categories (voiced unaspirated, voiceless unaspirated, voiceless aspirated) sharpened the non-English identification boundary on the VOT synthetic series for selected subjects. In a follow-up study, McClasky, Pisoni, and Carrell (1983) reported transfer of training in categorization of clear cases of labial stop consonants to identification of the nonnative voicing contrast in alveolar stops, and vice versa. However, neither of these studies investigated transfer of training effects to perception of natural speech.

Other perceptual training studies with the VOT dimension by Carney, Widin, and Viemeister (1977), Edman (1980), Samuel (1977), and Soli (1983) were motivated by questions concerning the limits of underlying psychophysical sensitivities rather than by a primary interest in language-specific differences in phonetic perception. These investigators employed fixed-standard discrimination tasks with immediate feedback and trained subjects for hundreds of trials over the course of several sessions. Results indicated large improvements in discriminability by AE speakers of stimuli with VOT values that encompass non-English phonetic contrasts. However, these investigators did not directly study possible transfer of training effects to the perception of foreign contrasts in natural speech.

The present study represents a merging of research on second-language learners' phonetic perception and psychophysical training studies. We were interested in whether we could modify the perception of AE word-initial /r/ and /l/ by adult Japanese learners of English, in the laboratory, using the psychophysical training task successfully employed by Carney et al. (1977). Questions of interest were: How easily does a change occur? What is the nature of the change? Does improvement in the training task with one set of stimuli transfer to other tasks and other stimuli? Does the change in perception of synthetic speech series transfer to tests of /r-l/ perception using real speech?

The /r-l/ contrast is not distinctive in Japanese phonology, and adult Japanese learners of English have a great deal of difficulty producing this contrast appropriately. They also have difficulties in perceptually differentiating these phonemes in AE natural speech minimal pairs (Mochizuki, 1981) which sometimes persist even after their production of the phonemes has become acceptable (Goto, 1971; Sheldon & Strange, 1982). Japanese adults show relatively poor identification and discrimination of this contrast in studies using a synthetic /r-l/ stimulus series (Miyawaki, Strange, Verbrugge, Liberman, Jenkins, & Fujimura, 1975; see also Mochizuki, 1981). MacKain et al., (1981) recently reported differences in identification and discrimination of a synthetic /r-l/ series that were correlated with the amount of conversational experience with native speakers. Japanese subjects with intensive conversational instruction and proportionally high everyday use of English produced very consistent identification functions and discrimination "peaks" at the phoneme boundary, indicating categorical perception. However, discrimination of cross-category pairs was still not as accurate as for native English speakers.

The design of the present study included comparisons of pretraining versus posttraining tests of natural speech minimal pairs contrasting /r/ and /l/ in several contexts, and categorical perception tests with two synthetic speech series contrasting /r/ and /l/ in word initial position. Training utilized one of the synthetic speech series and a fixed-standard AX discrimination procedure. Four subjects received training immediately after completing pretests, and four other subjects served as an independent control group. After these latter subjects were retested on pretest materials, they were given the same training experiences as the first four subjects.

With this design, several questions may be addressed regarding: (1) the nature of the improvement on the training task, (2) the effectiveness of training in producing native-like categorical perception of the trained-on stimuli and another synthetic speech series, and (3) transfer of training to perception of natural speech stimuli. Individual subjects' pretraining versus posttraining performance can be compared, as can group differences between trained and control subjects. Both types of comparisons are necessary for a clear interpretation of the results, inasmuch as we expected to find individual differences and because subjects were residing in the USA, and were receiving both informal and formal experience with English outside the laboratory.

METHOD

Subjects

Eight female native speakers of Japanese were recruited from a weekly intermediate-level English-as-a-second-language class sponsored by a foreign students organization of the University of Minnesota. The women were either students at the university or wives of Japanese exchange students, except S3 and S4, who were married to Americans. The subjects ranged in age from 25 to 33 years, and had resided in the USA 5 to 30 months at the time of testing. Their experience with and mastery of English varied considerably, but all reported that they had difficulty in perceiving and producing /ɪ/ and /l/ and all were eager to improve their English language skills. The subjects were divided into experimental and control groups for the initial training phase on the basis of convenience of scheduling.

Stimulus Materials

Three sets of stimulus materials were used in pretests and posttests of /r-l/ perception: (1) real-speech minimal pairs, (2) a "rock-lock" synthetic speech series, and (3) a "rake-lake" synthetic speech series. Each of these sets of materials have been utilized in previous studies and are described in detail in the original references.

Minimal pairs (Sheldon & Strange, 1982). Stimuli consisted of 16 pairs of words; four pairs contrasted /r/ and /l/ in each of four contexts–word-initial prevocalic, word-initial stop consonant + liquid clusters, word-medial intervocalic, and word-final postvocalic positions. In addition, eight minimal pairs contrasting other consonants and vowels served as fillers. An adult male native AE speaker produced each of the 48 words twice for a total of 96 utterances in all. Stimuli were recorded on audio tape (Revox A77 tape recorder and Electro-voice 660 microphone) with a 7-sec interstimulus interval. Native AE subjects identified all tokens unambiguously.

Rock-lock synthetic series (MacKain et al., 1981). A series of 10 stimuli was generated with the Haskins Laboratories OVE-IIIc synthesizer. The endpoints of the series simulated productions of "rock" (No. 1) and "lock" (No. 10) spoken by an adult male native speaker of AE. These endpoints differed in third formant (F3) onset frequency and transition, second formant (F2) onset and transition, and temporal pattern of the first formant (F1) initial steady-state and transition. Intermediate stimuli were generated by interpolating between endpoint parameter values in 10 nearly equal physical steps on F3 and F2 parameters, and in 5 equal steps on the F1 parameters. F3 onset varied from 1477 to 2594 Hz. F2 onset varied from 1067 to 1207 Hz. F1 steady-state duration ranged from 14 to 42 msec, changing in 7-msec steps on each two successive stimuli. Likewise, F1 transition duration ranged from 49 to 21 msec in five 7-msec steps. All 10 stimuli were 330 msec in duration and were identical over the last 220 msec. Intonation contours were identical for all 10 stimuli; amplitude contours varied somewhat on initial portions of the stimuli due to the differences in formant proximity which affect amplitude on cascade synthesizers such as the OVE-III. (See Appendix A for a detailed specification of the stimuli.)

Rake-lake synthetic series (Strange & Broen, 1981). A series of 10 stimuli was generated with the Haskins Laboratories parallel resonance synthesizer. The endpoints of the series simulated utterances of "rake" (No. 1) and "lake" (No. 10) spoken by an adult female native speaker of AE. Acoustic parameters that varied across the series were F3 onset and transition (from 1691 to 3196 Hz in 10 approximately equal steps), F2 onset and transition (from 1077 to 1382 Hz in 5 steps), and F1 onset and transition (from 255 to 432 Hz in 4 steps). Finally, the temporal pattern of F1 varied from a 36-msec steady-state and 66-msec transition for No.1 to a 66-msec steady-state and 18-msec transition for No. 10. Stimuli were 424 msec in duration; the final 256 msec of all 10 stimuli were identical. All stimuli had identical amplitude and intonation contours. (See Appendix C for a detailed specification of the 10 stimuli.)

Categorical perception tests. Identification and oddity discrimination tests were constructed separately for the rock-lock and rake-lake synthetic series. For identification tests of rock-lock, 20 repetitions of each of the 10 stimuli were randomized and recorded on audio tape in blocks of 20 stimuli, with a 3-sec interstimulus interval (ISI) and a 6-sec interblock interval (IBI). For oddity discrimination tests, the seven three-step comparison pairs (1-4, 2-5, . . . , 7-10) were arranged in triads in the six possible arrangements (AAB, ABA, BAA, ABB, BAB, BBA) for a total of 42 triads. Triads were randomized and recorded on audio tape with a 1-sec ISI and a 3-sec intertriad interval. Four such randomizations of the 42 triads were so recorded, for a total of 24 trials per comparison pair. Similar identification and three-step oddity discrimination tests were recorded on audio tape for the rake-lake stimuli. Test formats were the same except that the ISI was 2 sec and the IBI was 4 sec in the identification test.

Training stimuli. For use in the training task, the 10 rock-lock stimuli were filtered at 4750 Hz and converted to digital waveform files (10K/sec sampling rate with 12-bit resolution), using a PDP-8L laboratory computer. Software programs controlled the reconversion of these files to analog signals, which were filtered at 4750 Hz and presented via earphones to subjects seated in a one-person IAC acoustic chamber.

Procedures

Table 1 presents an outline of the design of the study. The sequence of tests for each subject is given from left to right.

Pretests. Subjects were tested in a quiet experimental room in groups of one to four in pretraining tests. Initial interviews, familiarization, and pretests were completed in two 1.5-h sessions. All subjects first completed the minimal pairs test, followed by identification of the rock-lock series and oddity discrimination of the rock-lock series. Seven of the eight subjects then completed identification of the rake-lake series, and oddity discrimination of the rake-lake series.[1] Instructions were given in both oral and written form, since the subjects' reading skills in English were very good. In the minimal-pairs test, the subjects responded after hearing each word by circling the appropriate member of the minimal pair printed on response forms. Prior to tests of each synthetic series, familiarization was given, in which the subjects heard 5 presentations of each of the endpoints, the series of 10 stimuli presented twice each in ascending, then descending order before identification tests, and 5 examples of oddity triads before discrimination tests. On identification tests, subjects responded by writing an "R" or "L" for each

Table 1
Sequence of Testing for Eight Japanese Subjects

	Pretest 1			Pretest 2			Training	Posttest		
Subjects	MP	R/Lock	R/Lake	R/Lock	R/Lake	MP	R Standard	R/Lock	R/Lake	MP
					E Group					
S1	X	X					4	X		X
S2	X	X	X				3	X	X	X
S3	X	X	X				3	X	X	X
S4	X	X	X				3	X	X	X
					C Group					
S5	X	X	X	X	X	X	4	X	X	X
S6	X	X	X	X	X	X	4	X	X	X
S7	X	X	X	X	X	X	3	X	X	X
S8	X	X	X	X	X	X	4	X	X	X

Note—MP, minimal-pairs test; R/Lock, identification and oddity discrimination tests of rock-lock stimulus series; R/Lake, identification and oddity discrimination tests of rake-lake stimulus series.

stimulus. On oddity discrimination tests, the subjects responded by marking whether the first, second, or third stimulus of the triad was the "different" one. Stimuli were presented binaurally via earphones (TDH-39) at 70 dB SPL.

Training. Each subject was tested individually during training sessions. The stimuli were presented binaurally over earphones (TDH-39) at 70 dB SPL. An all-step AX discrimination task with immediate feedback was used, in which the stimuli were presented in pairs, with a 1-sec ISI. The first stimulus was the standard and was fixed for a given block of trials. A block consisted of 18 trials, 9 same and 9 different. On different trials, the standard was paired with each of the 9 remaining stimuli of the series. On same trials, the standard was repeated. The sequence of same and different trials was selected randomly by the computer for each block. The subjects responded by pressing response buttons marked "S" and "D." Immediate feedback was given by lights that were illuminated over the correct response button. The next trial started 1 sec after the feedback light for the previous trial was terminated. All subjects completed training with two standards: Stimulus 8 (categorized by Americans as /l/) and Stimulus 3 or Stimulus 4 (both categorized consistently by Americans as /r/). (See Table 1 for assignment of /r/ standards to subjects.) The first session for each subject consisted of five consecutive blocks with each of the two standards. This session was considered training-task familiarization and the results were not analyzed further. Each subject then completed 14 to 18 training sessions, which took place on separate days over the course of about 3 weeks. A training session consisted of seven consecutive blocks of 18 trials with each standard.[2] The first block of trials with each standard was used as a warm-up and was not included in the data for the session. The order of standards was alternated from session to session for each subject and counterbalanced across subjects. The subjects were given feedback on their progress after each session.

Posttests. After the first four subjects completed training, all subjects were given the identification and discrimination tests of the rock-lock and rake-lake series and the minimal-pairs tests, using the same materials and procedures as for pretests. This constituted Pretest 2 for control subjects (S5-S8 in Table 1) and the posttest for trained subjects (S1-S4). Subjects S5-S8 were then given training sessions, using the same procedures, after which they were given the categorical perception tests and minimal-pairs test for the third time.

RESULTS

Pretests

Performance on pretraining tests indicated that all subjects had some difficulty in perceptually differentiating /r/ from /l/. The mean number of correct identifications on the minimal-pairs pretest was 44.25 (out of 64), or 69% correct. The subjects identified /r/-/l/ in word-final position best; the most errors were made on /r/-/l/ in consonant clusters, replicating the results of Sheldon and Strange (1982). The average number of correct identifications of word-initial /r/-/l/ was 10.25 (out of 16), or 64% correct. Individual subjects' scores ranged from 44% to 94% correct identification of initial /r/-/l/ words; only two of the eight subjects (S1, S4) identified these minimal pairs with above chance ac-

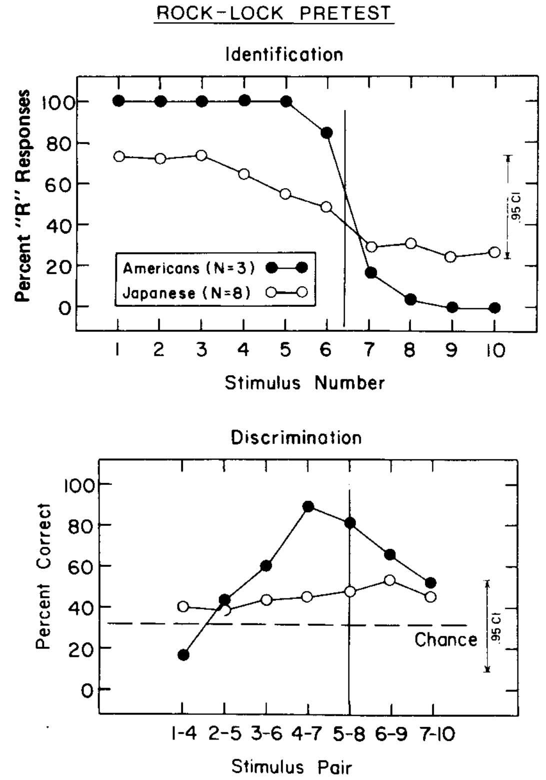


Figure 1. Pooled pretest identification functions (above) and oddity discrimination functions (below) for the rock-lock stimulus series.

curacy (by a binomial test of probabilities for a two-choice test, $p < .05$).

Figure 1 presents identification and oddity discrimination pretest results for the rock-lock synthetic series, pooled across all eight Japanese subjects (open circles). For comparison, the pooled functions of three native AE speakers, tested under identical conditions, are also given (closed circles). As is readily apparent, the Japanese subjects as a group were unable to identify the synthetic series consistently ($p > .05$) and average discrimination performance was not significantly above a chance level ($p > .05$ for all seven comparison pairs).[3] (See Appendix B for individual subjects' functions.)

Figure 2 presents pooled identification and oddity discrimination pretests for the rake-lake synthetic series. Again, it is readily apparent that the Japanese subjects neither identified nor discriminated this series the way Americans do. (See Appendix D for individual subjects' functions.)

Pretraining performance on the real-speech minimal-pairs test and identification tests of synthetic speech correlated well across subjects. Spearman rank-order correlations of minimal-pair scores with average identification scores for the four extreme rock-lock stimuli (1, 2, 9, and 10) showed a significant positive correlation ($r_S = +.89$, $p < .01$); minimal-pair scores also correlated significantly with identification of the four extreme rake-lake stimuli ($r_S = +.72$, $p < .05$). Thus, we can be confident that relative performance on these synthetic speech identification tasks accurately reflects perceptual difficulties foreign language learners have in differentiating this phoneme contrast.

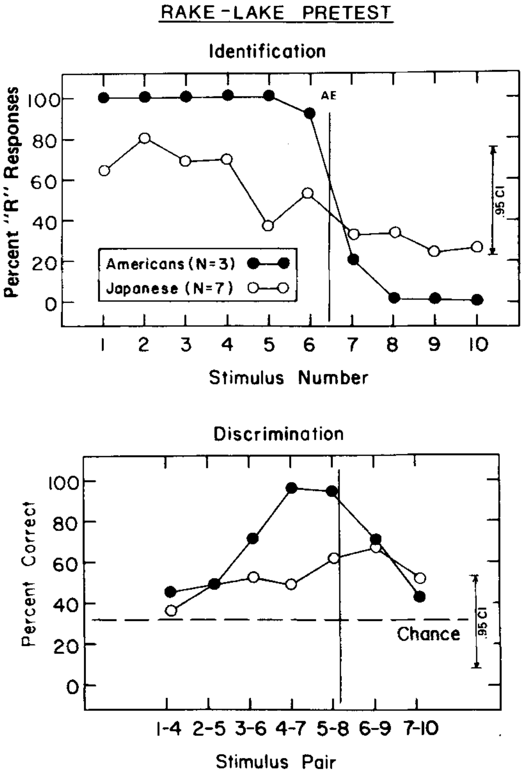


Figure 2. Pooled pretest identification functions (above) and oddity discrimination functions (below) for the rake-lake stimulus series.

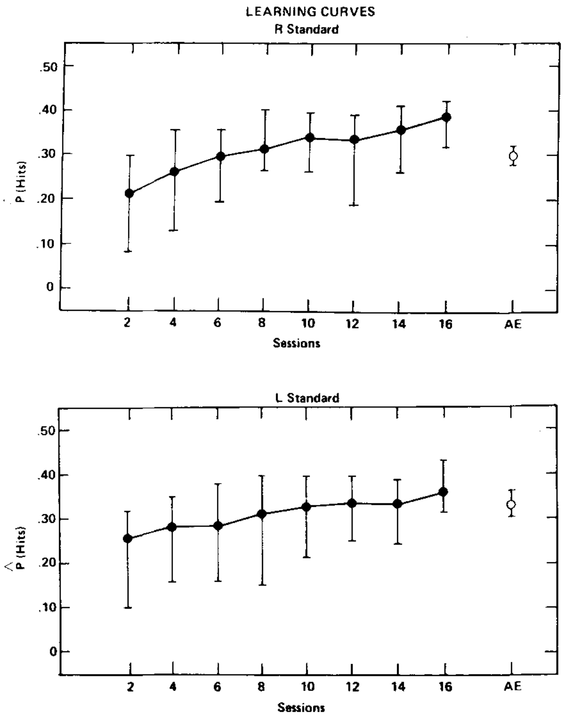


Figure 3. Performance in successive training sessions with the R standard (above) and the L standard (below).

Although there were considerable individual differences in performance on the pretests, the four subjects receiving training first (S1-24, henceforth labeled the E group) did not differ from the four control (C) subjects (S5-S8) as a group. However, this is because S1 and S4 were the best on pretest perception, whereas S2 and S3 were among the worst.

Training

Performance on the training task itself was characterized by gradual improvement over sessions with the greatest improvement in the first several sessions. Figure 3 presents the learning curves averaged across all eight subjects. The average probability of correct "D" responses, corrected for guessing [$\hat{P}$(Hits)] for successive blocks of training sessions, are given by the solid circles.[4] Ranges across subjects are indicated by vertical brackets. The open circles on the far right represent performance by two Americans on *initial* training trials; these thus serve as a reference for pretraining performance on this task by native AE speakers. Averaging over both standards, AE subjects' mean errors per block were 3.6 and $\hat{P}$(Hits) = .32. Japanese subjects as a group showed poorer discrimination than the Americans on initial trials [mean = 5.1 errors/block, $\hat{P}$(Hits) = .24], but, with training, they came to discriminate as well as or better than the Americans in their pretraining performance [mean = 2.9 errors/block, $\hat{P}$(Hits) = .35].

In order to illustrate more specifically the locus of this improvement, Figure 4 presents the pooled all-step AX discrimination functions for the first three training sessions (open circles) versus the last three training sessions (closed circles) for the four subjects who were trained on Standard 3 (above) and Standard 8 (below). Figure 5 presents the functions for the subjects who were trained on Standards 4 and 8. The results are plotted as the percent hits ("D" responses on D trials) for each of the nine comparison stimuli. The squares above the standards give the false-alarm rates (percent "D" responses on S trials).

Recall that AE subjects identified Stimuli 1-5 "R" on 100% of the trials and Stimuli 8-10 "L" 100%, but categorized Stimuli 6 and 7 inconsistently. Looking first at the data for the initial three sessions, it is apparent that discrimination even of stimuli from different AE phoneme categories was not perfect for Japanese subjects, although it was better than would be predicted from their identification and oddity discrimination pretests. This no doubt reflects the fact that decreased stimulus uncertainty and memory load makes the fixed-standard AX task easier. Discrimination of within-category stimuli was not better than chance; the proportion of "D" responses on D trials was about equal to "D" responses on S trials.

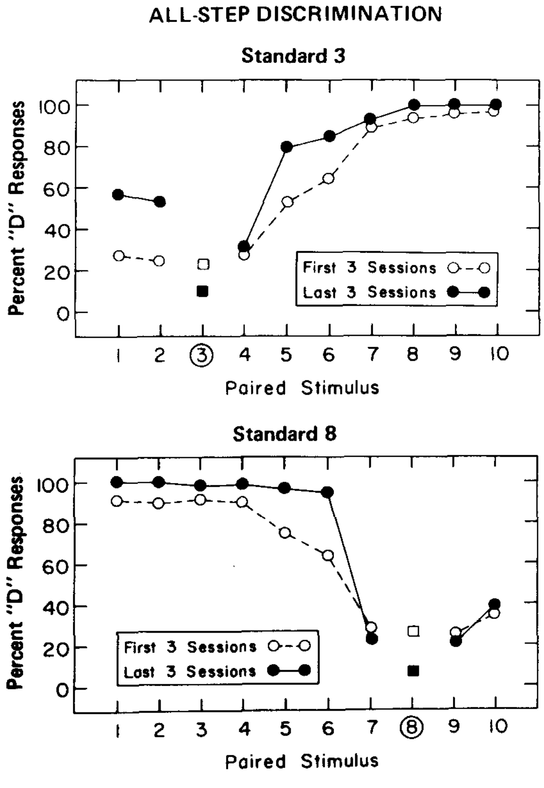


Figure 4. All-step discrimination functions of initial and final training sessions, pooled across the four subjects trained on R standard 3.

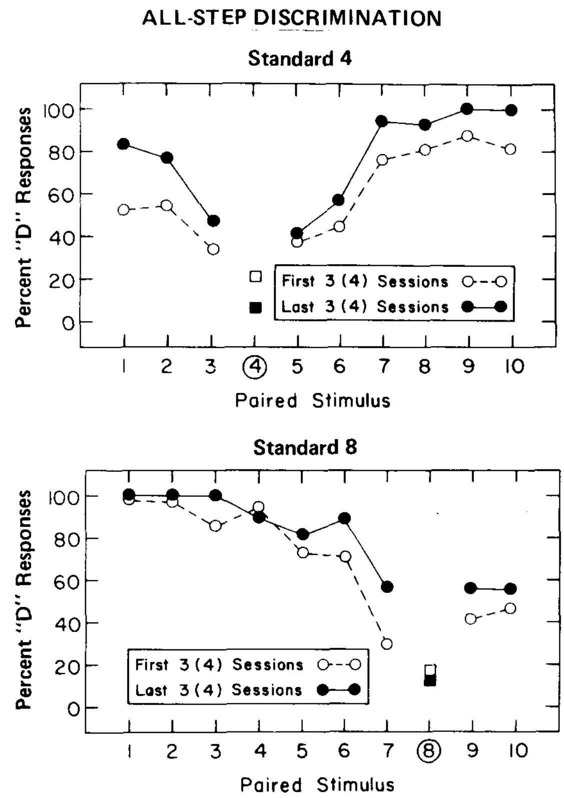


Figure 5. All-step discrimination functions of initial and final training sessions, pooled across the four subjects trained on R standard 4.

By the final three sessions, as Figure 4 indicates, cross-category stimuli were discriminated almost perfectly from each standard for these four subjects. Discrimination improved most for Stimuli 5 and 6 versus both standards, and for within-/r/-category Stimuli 1 and 2 versus Stimulus 3. In contrast, within-/l/-category discrimination showed little improvement. Figure 5 shows a similar pattern for the other four subjects. Cross-category stimuli were discriminated almost perfectly, except for Stimuli 4 and 5 versus Stimulus 8. Discrimination of ambiguous Stimuli 6 and 7 versus each standard showed considerable improvement, as did within-/r/-category discrimination (Stimuli 1 and 2 vs. Stimulus 4). Again, discrimination within the /l/ category showed the least improvement. These pooled functions are fairly representative of individual performance. Although subjects showed considerable variability, all eight subjects improved with training.

Posttests

For each of the posttests, two kinds of comparisons are possible: (1) performance by E versus C subjects, and (2) pretraining versus posttraining performance by all eight subjects, both individually and as a group. In the following discussion, the independent groups comparison is considered first, followed by the repeated measures comparisons. Performance on each of the three sets of stimuli is considered separately.

Rock-lock series. Figure 6 presents the pooled identification and oddity discrimination functions for the four E subjects on posttests, compared with the pooled functions for the four C subjects on their second pretest. This comparison allows us to partial out specific effects of training with the rock-lock stimuli from improvement due to repeated testing or extra-experimental experiences. As the functions show, training had a clearly positive effect on subsequent categorical perception tasks with the trained-on stimuli, over and above improvement due to task familiarity or other factors.

Identification by the E group was marked by relatively consistent labeling of within-category stimuli, and a fairly abrupt crossover between categories occurring between Stimuli 6 and 7, as was the case for AE subjects. In contrast, the pooled identification function for the C group was not significantly changed from their first pretest identification performance.

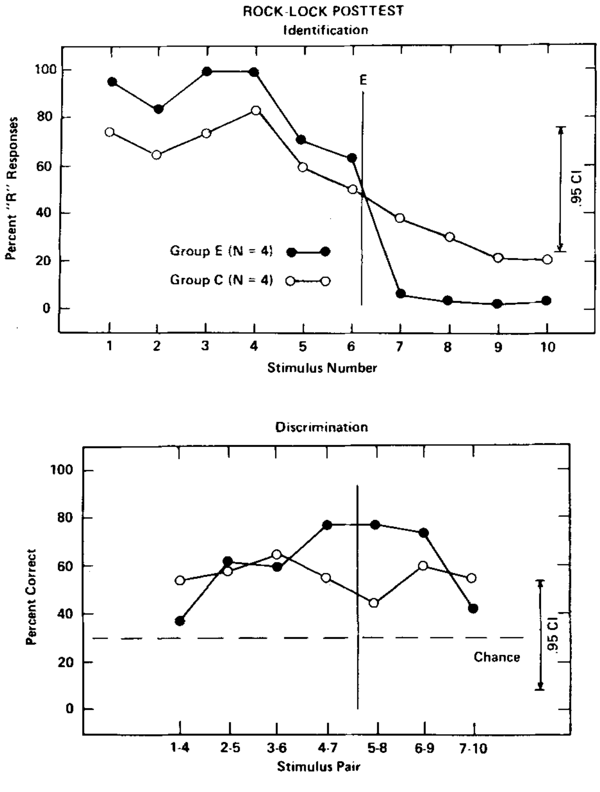


Figure 6. Pooled identification functions (above) and oddity discrimination functions (below) for experimental subjects on posttraining tests of the rock-lock series (solid circles), compared with control subjects on the second pretraining tests (open circles) of the rock-lock series.

Comparison of pooled oddity discrimination performance of E versus C groups also revealed obvious differences as a function of the training experience. The C group did improve from Pretest 1 to Pretest 2 in average discrimination accuracy. The group function indicates that all pairs, except 5-8, were discriminated above chance on Pretest 2 ($p < .05$), whereas no pairs were discriminated above chance on Pretest 1 ($p > .05$). However, note that the "peak" of most accurate discrimination (3-6) did not occur in the region of the AE phoneme boundary (or the E group's identification boundary). In contrast, the E group function showed a single broad peak of most accurate discrimination (averaging 76%) for pairs which crossed their own and the AE identification boundary; within-category pairs were discriminated relatively less accurately. However, average discrimination (61%) on Stimuli 2-5 and 3-6 was also above chance ($p < .01$).

A comparison of pretraining versus posttraining categorical perception tests for each of the eight subjects on the rock-lock series revealed that seven of the eight subjects improved as a function of the training. Only S5 did not; this subject made errors on cross-category comparisons even in her final training sessions. Improvement in identification was analyzed statistically for each subject by repeated measures t tests, using stimuli as the sampling variable. Results are given in Table 2. Although S2, S3, and S6 showed significant improvement over pretest identification, they still produced nonmonotonic functions at posttest. S1, S4, S7, and S8 produced highly consistent and discontinuous functions similar to those of native AE speakers.

Pre- versus posttraining comparisons of oddity discrimination revealed similar results. The pre- versus posttraining performance of each subject was assessed by chi-square tests of the three comparison pairs that straddled the subject's identification boundary. The results are reported in Table 2. Again, S5 showed no improvement from Pretest 2 to posttest; S4 and S7 also failed to show significant improvement over pretests. However, these two subjects performed relatively well at pretest and posttest discrimination of cross-category pairs with over 75% correct. The other subjects showed marked improvement for pairs that crossed their identification category boundary; within-category discrimination also improved for some subjects.

Identification posttest functions for the seven subjects who improved with training were characterized by relatively abrupt crossovers between categories. However, these category boundaries varied considerably among the seven subjects; the 50% crossover in unsmoothed functions was between Stimuli 4 and 5 for three subjects and between Stimuli 6 and 7 for four subjects.[5] Thus, pooled identification and oddity discrimination functions, plotted by stimulus number as in Figures 1 and 6, are not representative of individuals' performance. However, by taking each subject's identification boundary as a reference point and adjusting

Table 2
Statistical Analyses of Improvement from Pre- to Posttraining Categorical Perception Tests of the Rock-Lock Series by Individual Japanese Subjects

	Identification		Discrimination	
Subjects	t(9df)*	p value†	χ^2(3df)**	p value
S1	4.070	<.002	17.72	<.001
S2	5.170	<.001	9.26	<.05
S3	5.412	<.001	20.12	<.001
S4	3.788	<.01	6.51	n.s.
S5	−2.414	n.s.	1.20	n.s.
S6	3.869	<.002	28.84	<.001
S7	4.636	<.001	3.13	n.s.
S8	4.275	<.002	13.14	<.01

*Note—For S5-S8, Pretest-2 data were compared with posttests. *Repeated measures t tests were performed on pre- vs. posttraining number correct, using stimuli as the sampling unit. "Correct" was defined with reference to each subject's posttest category boundary (50% crossover in unsmoothed functions) or the AE boundary for subjects with no clear posttest boundary. **Chi-square tests compared pre- vs. posttraining frequencies of correct and incorrect responses on each of the three comparison pairs straddling each subject's posttest identification boundary, as defined in *. †One-tailed test.*

data points accordingly, group identification and oddity discrimination functions that are more representative can be presented. Figure 7 presents adjusted pre- and posttraining data pooled over the seven subjects who produced clear identification boundaries on posttests. Pre- and posttraining oddity discrimination functions were also adjusted according to each subject's posttest identification boundary (C indicates cross-category pairs; W represents within-category pairs).

As with the E versus C group comparison shown above, a change toward categorical perception of the rock-lock series, as a function of training, is clearly indicated. However, group identification, even when adjusted for boundary differences, was less consistent than that of native AE subjects, and discrimination of cross-category pairs was not as accurate (see Figure 1 for AE functions). We can thus conclude that training with a fixed-standard AX discrimination task resulted in improved (categorical) perception of the training stimuli, as tested by the (more demanding) identification and oddity discrimination tasks. However, performance was still not completely native-like for all subjects, even after thousands of trials with the specific stimuli being tested.

Rake-lake series. Figure 8 compares the group identification and oddity discrimination functions for the E group (N = 3) on the rake-lake posttest with those for the C group (N = 4) on rake-lake Pretest 2. The differences in performance shown here reflect a transfer of the effects of training with the rock-lock stimuli to perception of another synthetic series with quite different acoustic properties.

As the functions show, E group identification was more consistent than C group identification, and discrimination was somewhat more accurate especially for cross-category pair 5-8. However, the differences were not as great as for the rock-lock comparison, especially on discrimination. In part, this is because C group subjects improved considerably from Pretest 1 to Pretest 2 in discrimination of all seven pairs. On both pretests, there were two peaks of relatively accurate discrimination. This was due to the tendency of some subjects to categorize stimuli at both ends of the series as "L," while labeling the middle stimuli as "R." (See individual functions in Appendix D.)

A comparison of pre- versus posttraining categorical perception tests of the rake-lake series for individual subjects indicated that five of the seven subjects tested improved in identification or discrimination as a function of training (see Appendix D). Improvement in identification by each subject was analyzed statistically by repeated measures t tests; the results are given in Table 3. Although S3, S5, S6, and S8 showed significant improvement over pretest performance levels, only S3 and S4 produced discontinuous posttest functions resembling those of native AE speakers. S5 (who did not improve on rock-lock tests) and S2 both produced very inconsistent functions. The other three subjects (S6, S7, S8) produced above-chance identification of the four end stimuli at posttest, but the functions were more con-

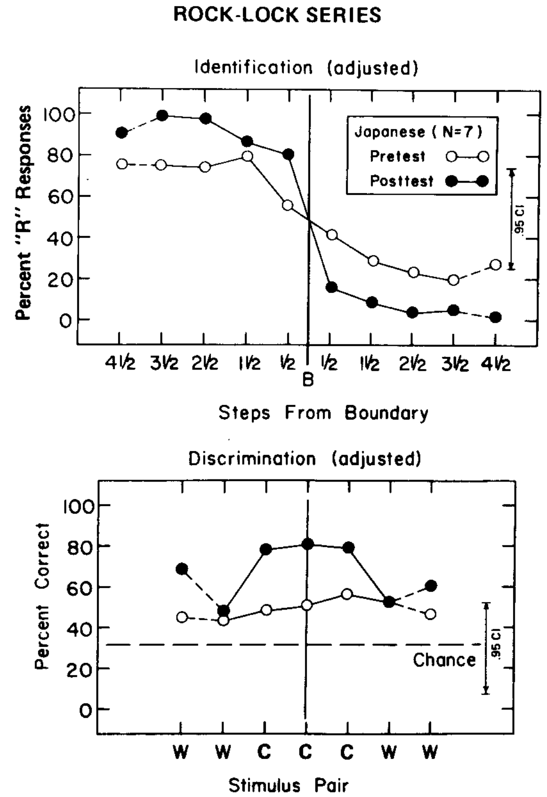
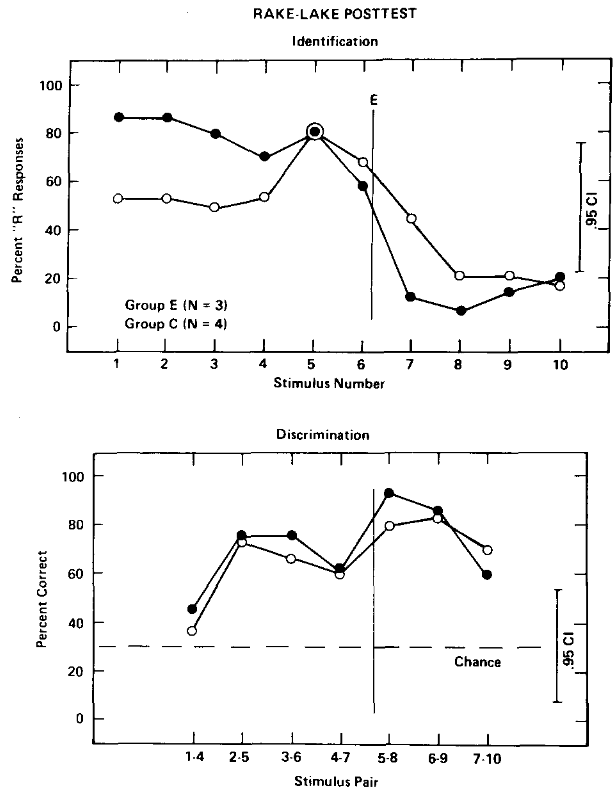


Figure 7. Comparison of pre- versus posttraining adjusted identification (above) and oddity discrimination tests (below) on the rock-lock series, pooled across seven subjects.


RAKE-LAKE POSTTEST

Identification

Discrimination

Figure 8. Pooled identification functions (above) and oddity discrimination functions (below) for experimental subjects on posttraining tests of the rake-lake series (solid circles), compared with control subjects on the second pretraining tests (open circles) of the rake-lake series.

tinuous than those of AE subjects and some showed nonmonotonicities.

Pre- versus posttraining discrimination of each subject's three cross-category pairs was analyzed by chi-square tests; the results are given in Table 3. Only S2 improved significantly over pretest performance; however, other subjects produced cross-category discrimination well above chance on both pretests and posttests. In addition, discrimination was above chance even for within-category comparison pairs on posttests for most subjects. Many functions contained two peaks of relatively accurate discrimination.

Table 3
Statistical Analyses of Improvement from Pre- to Posttraining Categorical Perception Tests of the Rake-Lake Series by Individual Japanese Subjects

	Identification		Discrimination	
Subjects	t(9df)	p value*	χ^2(3df)	p value
S1				
S2	1.70	n.s.	16.88	<.001
S3	3.03	<.01	6.94	n.s.
S4	1.34	n.s.	4.56	n.s.
S5	3.23	<.01	5.38	n.s.
S6	8.39	<.001	1.66	n.s.
S7	1.20	n.s.	1.36	n.s.
S8	2.05	<.05	3.10	n.s.

Note–For S5-S8, Pretest-2 data were compared with posttests.
**One-tailed test.*

As with the rock-lock tests, the five subjects (S3, S4, S6, S7, S8) who had clearly defined phoneme category boundaries on identification posttests varied in their placement of the boundary between rake and lake. Thus, pre- and posttraining identification and discrimination data for these five subjects were adjusted by reference to their posttest boundaries and pooled, as shown in Figure 9. Again, some improvement as a function of training can be seen. However, identification was not as consistent as for AE subjects (see Figure 2), nor was it as consistent as these subjects' posttest identification of the rock-lock series. Cross-category discrimination was about 80% accurate, compared with greater than 90% accuracy for AE subjects (see Figure 2). However, this high level of discrimination cannot be attributed to the training experience per se, because performance on Prestest 2 by subjects S5-S8 was as accurate as posttest performance. Thus, transfer of training to an acoustically different stimulus series was less than complete even for these five subjects.

Minimal pairs. Table 4 presents pre- and posttraining performance on the initial /r/-/l/ minimal pairs by E group and C group subjects. On the average, there was no significant improvement in identification of these real-speech stimuli as a function of training [$t(3) = 1.44$, $p > .05$]. A repeated measures t test of all eight subjects' pre- versus posttraining identification also revealed no significant overall improvement [$t(7) = 1.71, p > .05$]. However, four subjects (S1, S4, S5, S6) identified these words above a chance level ($p < .05$ by a binomial test) at posttest, whereas only S1 and S4 performed above chance at pretest. Note, however, that S5 was the subject who showed little improvement on the rock-lock and rake-lake categorical perception tests! Only S8 improved from below-chance to above-chance performance on minimal pairs contrasting /r/ and /l/ in initial consonant clusters. Thus, we can conclude that transfer of training with the rock-lock synthetic series to perception of real-speech contrasts of /r/ and /l/ was minimal.

DISCUSSION

For purposes of summarizing the results of this study, it is informative to classify subjects' performance on the several tests according to descriptive criteria that capture the nature of their perceptual differentiation of /r/ and /l/. Classification of performance on minimal pairs tests was straightforward, and is reported in terms of binomial probabilities ($p = .50$). Perceptual performance on tests of synthetic series was classified as chance, inconsistent, continuous, or categorical, according to the following criteria. *Chance*

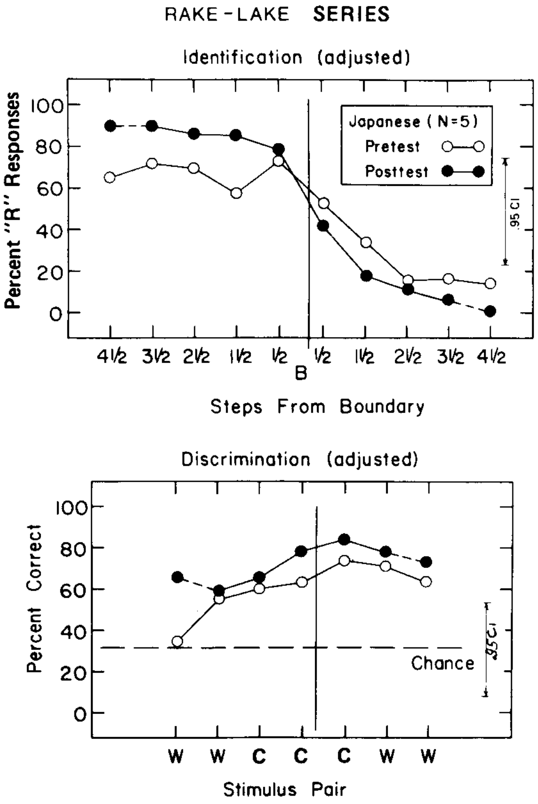


Figure 9. Comparison of pre- versus posttraining adjusted identification (above) and oddity discrimination tests (below) on the rake-lake series, pooled across five subjects.

performance consisted of an identification function in which all 10 stimuli were labeled with less than 80% accuracy and discrimination was not above 58% correct for any comparison pair. (This constitutes the adoption of a 99% confidence level for both tests.) *Inconsistent* perception designated those identification and discrimination functions for which performance was above the chance level ($p < .01$) for some stimuli, but the identification function was nonmonotonic and discrimination peaks were not predictable from identification category boundaries. *Continuous* perception consisted of an identification function with no major nonmonotonicity (i.e., no greater than a 25% reversal) and three of the four end stimuli identified with at least 80% accuracy, coupled with a discrimination function showing above-chance performance for cross-category or within-category comparison pairs, or both. Perception was classified as *categorical* if identification functions showed highly consistent labeling (>90%) of all but the two stimuli closest to the category boundary and discrimination functions had peaks of most accurate discrimination for cross-category comparison pairs. These classifications are thus ordered in terms of increasing perceptual differentiation and similarity to native speakers' performance. Table 5 summarizes the results for each subject according to these classifications.

Table 4
Correct Identification of Initial /r/-/l/ Minimal Pairs on Pre- and Posttraining Tests

	Pretest 1		Pretest 2		Posttest	
	Mean	Range	Mean	Range	Mean	Range
Group E	11.0	7-15			11.25	7-14
Group C	9.5	8-11	8.5	6-10	11.0	10-12
Overall	10.25*				11.12	

Note—*Total identifications possible = 16.* **Average over Group E Pretest 1 scores and Group C Pretest 2 scores.*

On the basis of this classificatory scheme, one can readily see the positive effects of training on subsequent tests of perception of synthetic series contrasting /r/ and /l/. All subjects changed toward better perceptual differentiation, regardless of their pretraining performance levels. Subjects S1 and S4 were the best perceivers at pretest and could differentiate real-speech pairs above chance accuracy. Nevertheless, the laboratory training experience served to improve their performance on synthetic series such that they resembled native AE speakers at posttest. Indeed, S4 categorically perceived both the trained-on series and the transfer rake-lake series after training. S7 and S8, who showed continuous or inconsistent perception of the rock-lock series at pretests, also improved with training; at posttest, their perception of the trained-on stimuli was categorical. Perception of the transfer series also improved for these subjects. Subjects S3 and S6 also showed great gains with training, going from chance or inconsistent performance to continuous perception of both rock-lock and rake-lake series as a function of training. Subjects S2 and S5 were among the poorest perceivers at pretest and showed little overall improvement with training. Their posttest performance remained inconsistent for both trained-on and transfer stimulus materials.

Returning to the questions posed in the introduction, we may conclude on the basis of these data that training with a fixed-standard AX discrimination task improved subjects' perceptual differentiation of synthetic instances of word-initial /r/ and /l/. Increases in discrimination accuracy during the training task were gradual and subjects reported that the task remained quite difficult, requiring close attention. Improvement in AX discrimination of the rock-lock stimuli carried over to categorical perception tests of these same stimuli. The oddity discrimination task has more stimulus uncertainty and is more demanding in terms of memory load, since it requires the comparison of three stimuli separated by 1-sec intervals. The identification task requires categorization of stimuli on the basis of internalized representations of the phoneme categories. In neither task did the subjects receive feedback about the accuracy of their responses. Thus, it is not trivial that most subjects were able to perform better on these tests after training.

Table 5
Summary of Results of Training for Individual Subjects

	Pretest 1			Pretest 2			Posttest		
Subjects	MP	R/Lock	R/Lake	R/Lock	R/Lake	MP	R/Lock	R/Lake	MP
1	<.01	Cont					Cat		<.01
2	n.s.	Ch	Inc				Inc	Inc	n.s.
3	n.s.	Inc	Inc				Cont	Cont	n.s.
4	<.01	Cont	Cont				Cat	Cat	<.01
5	n.s.	Ch	Inc	Inc	Inc	n.s.	Inc	Inc	<.05
6	n.s.	Ch	Inc	Ch	Inc	n.s.	Cont	Cont	<.05
7	n.s.	Cont	Inc	Cont	Cont	n.s.	Cat	Cont	n.s.
8	n.s.	Inc	Inc	Inc	Inc	n.s.	Cat	Cont	n.s.

Note–MP = minimal pairs test; R/Lock = categorical perception tests of rock-lock stimuli; R/Lake = categorical perception tests of rake-lake stimuli. Ch = chance perception (see text); Inc = inconsistent perception (see text); Cont = continuous perception (see text); Cat = categorical perception (see text).

There appeared to be no correlation between which /r/ standard was presented in training and the placement of the category boundary on subsequent rock-lock identification tests. Variability in boundary placement was due to subjects' differing in their categorization of Stimuli 5 and 6 as either primarily "R" or primarily "L." Recall that these were stimuli on which great improvement occurred during training with both /r/ and /l/ standards. We can speculate that subjects may actually have differentiated three categories during identification posttests, but were unable to report this because of the restriction to two response labels. This is borne out in part by the tendency for posttest discrimination of within-category pairs to be more accurate than would be predicted from the two-category labeling data. It would be of interest in future studies to test subjects after training in an open-response-set identification task.

With respect to questions concerning the transfer of training to a new /r/-/l/ synthetic series with different acoustic properties, the answer can only be a tentative "Yes" on the basis of these data. As Table 5 indicates, five subjects who showed categorical or continuous perception on rock-lock posttests also performed relatively well on tests of the rake-lake series after training. However, transfer was less than complete in that posttest functions were less consistent than those for the trained-on stimuli. Only S4 produced functions that were classified as categorical (although we can speculate that S1 may have produced categorical perception of this series had she been tested). Also, there were indications that some of the improvement from pre- to posttraining performance might well be attributable to extraexperimental experience or to familiarity with the stimuli and procedures. This can be surmised from the improvement from Pretest 1 to Pretest 2 of some subjects on discrimination (but not identification) tests and also in the better discrimination performance at Pretest 1 on rake-lake tests (which were conducted second) over rock-lock tests.

Finally, the results of minimal-pairs tests showed positive transfer to word-initial /r/ versus /l/ for only two subjects, including S5, who had performed most poorly on training and posttests of synthetic series. Thus, we cannot conclude that this training experience generalized to perception of the phoneme contrast in real speech by a native AE speaker. However, this lack of generalization comes as no surprise when one considers the specificity of the training task and materials employed. Future studies should be designed in such a way that subjects learn to abstract the relevant parameters which differentiate the phonemes while ignoring the acoustic and phonetic contextual variations that are not distinctive with respect to the contrast. This would include training of the contrast with more than one set of stimuli and in more than one phonetic context.

As reviewed briefly in the introduction, previous training studies employing synthetic VOT series have yielded generally different results from those reported here. Pisoni et al. (1982) reported rapid improvement with relatively little training in differentiating "prevoiced" from "voiced" initial stop consonants, at least for some subjects. Edman (1980, see also Soli, 1983) reported rapid, "quantal" improvement in discrimination of VOT when immediate feedback was instituted. These results contrast with the gradual improvement in performance reported here. In addition, transfer of training to new (synthetic) VOT series appears to have been more complete than the transfer shown here (Edman, 1980; McClaskey et al., 1983).

There are several possible reasons why the modification of perception of the /r/-/l/ contrast by Japanese adults may be more difficult than modification of VOT perception by Americans. These include phonological, phonetic, and acoustic differences between the two types of contrasts. The Japanese were required to learn a distinction between two phonemes, neither of which are similar phonetically to any phoneme in their language. In Japanese, the phoneme /r/ is realized phonetically as an apical flap [ɾ] (Price, 1981), and is minimally contrasted with the oral stop [d]. Thus, neither AE approximant [ɹ] nor [ℓ] is phonetically similar to any phoneme in Japanese. In the case of the studies with VOT, Americans were required to learn to discriminate between voiced and voiceless, unaspirated variants of initial stop consonants, which in English are allophonic variants of (phonologically) voiced stop consonants. That is, two of the three phoneme categories encompassed by the VOT synthetic series represent a native contrast in English, and the third category includes speech sounds that are often produced in English, as free variants in word-initial contexts and as allophonic variants in intervocalic phonetic contexts. Thus, AE subjects have had considerable experience with the phonetic categories being trained, although not as distinctive (functional) categories in their native language.

There is some evidence that learning a nonnative contrast that includes one or more phonetic variants that are similar to those present in the native language may be easier than learning a contrast in which both categories are phonetically dissimilar to any native phoneme. Best, MacKain, and Strange (1982) tested Japanese learners of English on three synthetic series contrasting /w/-/y/, /w/-/r/, and /r/-/l/. The first contrast occurs in Japanese, although phonetically, /w/ is unrounded. The second contrast thus represents a distinction between /w/, which occurs in Japanese, and approximant [ɹ], which does not occur. As predicted, identification and discrimination by Japanese of the /w/-/y/ series was highly similar to the performance of American controls. Performance on /w/-/r/, although not as good as for Americans, was better than performance on the /r/-/l/ series. Phoneme category boundaries were steeper and cross-category discrimination was more accurate for /w/-/r/ than for /r/-/l/.

A second possible reason for the discrepancy in results of training studies with VOT and /r/-/l/ may lie in the "intrinsic difficulty" of the phonetic variation. Approximant [ɹ] is a relatively rare variant of the phoneme /r/ in the languages of the world, especially as contrasted with approximant /l/. In contrast, voicing distinctions between voiced and voiceless unaspirated phonetic variants are very common. We may speculate that there is a correlation between ease of perceptual differentiability and usage in languages.

Finally, consideration of the acoustic specification of the two types of contrasts leads to a prediction that perception of voicing distinctions may be less difficult to modify. The acoustic parameter involved in training studies with Americans was the duration of "prevoicing," that is, the extent of low-frequency periodic energy preceding the onset of the stop burst and upper formant energy. Thus, the differentiating parameter involved a temporally distinct acoustic component. In the case of the /r/-/l/ distinction, the differentiating spectral and temporal parameters occur within the context of a complex signal in which there are simultaneously occurring spectral components that are not distinctive with respect to the contrast. It is thus reasonable to assume that subjects would have more difficulty focusing their attention on the differentiating aspects of the signals. Indeed, it is not possible to "hear" individual formants as separable events, so subjects could not perceptually separate differential from nondifferential acoustic components. (See also Edman, 1980, Experiment 2.)

Further research is necessary with different phonetic contrasts, different stimulus materials, and subjects from different language backgrounds before we can make any general statement about the efficacy of laboratory training techniques in modifying the phonetic perception of adults learning foreign languages. Likewise, general claims about the malleability of phonetic perceptual processes in adulthood cannot be made on the basis of one or two specific instances.

The accumulating evidence on the development of speech perceptual processes suggests that human infants are equipped with rather remarkable capacities to discriminate phonetically relevant acoustic parameters of speech (Aslin, Pisoni, & Jusczyk, 1983). In the course of learning their native language, children come to recognize and categorize speech utterances on the basis of language-specific distinctive phonetic contrasts. We conceive of this as the development of selective perceptual skills that become automatic, such that phonologically relevant phonetic parameters are abstracted and differentiated, whereas nondistinctive variations are filtered out and ignored (Strange & Broen, 1980; see also Aslin & Pisoni, 1980). Second-language learning, then, involves the reeducation of selective perceptual processes. The research presented here and elsewhere suggests that this reeducation process requires intensive instruction and considerable time and effort, at least for some types of phonetic contrasts. We do not yet know which methods will best produce generalization of training and relatively automatic phonetic perception. Research on this problem should shed light on the basic processes involved in speech perception and the role of linguistic experience in the development and modification of those processes.

REFERENCES

Abramson, A. S., & Lisker, L. (1967). Discriminability along the voicing continuum: Cross-language tests. *Proceedings of the 6th International Congress of Phonetic Sciences*, 569-573.

Aslin, R. N., & Pisoni, D. B. (1980). Some developmental processes in speech perception. In G. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology: Vol. 2. Perception* (4th ed.). New York: Academic Press.

Aslin, R. N., Pisoni, D. B., & Jusczyk, P. W. (1983). Auditory development and speech perception in infancy. In P. Mussen (Ed.), *Carmichael's manual of child psychology: Vol. 2. Infancy and the biology of development* (4th ed., pp. 573-687; M. M. Haith & J. J. Campos, Vol. 2 Eds.). New York: Wiley.

Beddor, P. S., & Strange, W. (1982). Cross-language study of the perception of the oral-nasal distinction. *Journal of the Acoustical Society of America*, 71, 1551-1561.

Best, C. T., MacKain, K. S., & Strange, W. (1982). A cross-language study of categorical perception for semi-vowel and liquid glide contrasts. *Journal of the Acoustical Society of America*, **71**, Supplement 1, S76. (Abstract)

Briere, E. (1966). An investigation of phonological interference. *Language*, **44**, 786-796.

Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America*, **54**, 421-428.

Carney, A. E., Widin, G. P., & Viemeister, N. F. (1977). Noncategorical perception of stop consonants differing in VOT. *Journal of the Acoustical Society of America*, **62**, 961-970.

Edman, T. R. (1980). *Learning of intraphonemic discrimination for several synthetic speech continua*. Unpublished doctoral dissertation, University of Minnesota.

Fledge, J. E. (in press). The effect of linguistic experience on Arabs' perception of the English /s/ vs /z/ contrast. *Folia Linguistica*, **18**.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds "l" and "r." *Neuropsychologia*, **9**, 317-323.

Gottfried, T. L. (1984). Effects of consonant context on the perception of French vowels. *Journal of Phonetics*, **12**, 91-114.

Keating, P. A., Mikos, M. J., & Ganong, W. F. (1981). A cross-language study of range of voice onset time in the perception of initial stop voicing. *Journal of the Acoustical Society of America*, **70**, 1261-1271.

Lisker, L. (1970). *On learning a new contrast* (Status Report on Speech Research, SR-24, pp. 1-15). New Haven, Conn: Haskins Laboratories.

MacKain, K. S., Best, C. T., & Strange, W. (1981). Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics*, **2**, 369-390.

Marckwardt, A. H. (1944). An experiment in aural perception. *The English Journal*, **33**, 212-214.

Marckwardt, A. H. (1946). Phonemic structure and aural perception. *American Speech*, **21**, 106-111.

McClasky, C. L., Pisoni, D. B., & Carrell, T. D. (1983). Transfer of training of a new linguistic contrast in voicing. *Perception & Psychophysics*, **34**, 323-330.

Miyawaki, K., Strange, W., Verbrugge, R. R., Liberman, A. M., Jenkins, J. J., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of /r/ and /l/ by native speakers of Japanese and English. *Perception & Psychophysics*, **18**, 331-340.

Mochizuki, M. (1981). The identification of /r/ and /l/ in natural and synthesized speech. *Journal of Phonetics*, **9**, 283-303.

Nemser, W. (1971). *An experimental study of phonological interference in the English of Hungarians*. Bloomington: Indiana University Press.

Oller, D. K., & Eilers, R. E. (1983). Speech identification in Spanish- and English-learning 2-year-olds. *Journal of Speech and Hearing Research*, **26**, 50-53.

Pisoni, D. B. (1971). *On the nature of categorical perception of speech sounds*. Unpublished doctoral dissertation, University of Michigan.

Pisoni, D. B., Aslin, R. N., Perey, A. J., & Hennessy, B. L. (1982). Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology: Human Perception and Performances*, **8**, 297-314.

Price, P. J. (1981). *A cross-linguistic study of flaps in Japanese and in American English*. Unpublished doctoral dissertation, University of Pennsylvania.

Samuel, A. G. (1977). The effect of discrimination training on speech perception: Noncategorical perception. *Perception & Psychophysics*, **22**, 321-330.

Sapon, S. M., & Carroll, J. B. (1958). Discrimination perception of speech sounds as a function of native language. *General Linguistics*, **3**, 62-72.

Sheldon, A., & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics*, **3**, 243-261.

Singh, S., & Black, J. W. (1966). Study of twenty-six intervocalic consonants as spoken and recognized by four language groups. *Journal of the Acoustical Society of America*, **20**, 372-387.

Soli, S. D. (1983). The role of spectral cues in discrimination of voice onset time differences. *Journal of the Acoustical Society of America*, **73**, 2150-2165.

Stevens, K. N., Liberman, A. M., Studdert-Kennedy, M., & Öhman, S. E. G. (1969). Cross language study of vowel perception. *Language and Speech*, **12**, 1-23.

Strange, W. (1972). *The effects of training on the perception of synthetic speech sounds: Voice onset time*. Unpublished doctoral dissertation, University of Minnesota.

Strange, W., & Broen, P. A. (1980). Perception and production of approximant consonants by three-year-olds: A first study. In G. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology: Vol. 2. Perception* (pp. 117-154). New York: Academic Press.

Strange, W., & Broen, P. A. (1981). The relationship between perception and production of /w/, /r/, and /l/ by three-year-old children. *Journal of Experimental Child Psychology*, **31**, 81-102.

Streeter, L. A., & Landauer, T. K. (1976). Effects of learning English as a second language on the acquisition of a new phonemic contrast. *Journal of the Acoustical Society of America*, **59**, 448-451.

Trehub, S. E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Development*, **47**, 466-472.

Werker, J. F. (1982). *The development of cross-language speech perception: The effect of age, experience, and context on perceptual organization*. Unpublished doctoral dissertation, University of British Columbia.

Werker, J. F., Gilbert, J. H. V., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, **52**, 349-355.

Williams, L. (1977). The perception of stop consonant voicing by Spanish-English bilinguals. *Perception & Psychophysics*, **21**, 289-297.

Williams, L. (1979). The modification of speech perception and production in second-language learning. *Perception & Psychophysics*, **26**, 95-104.

NOTES

1. S1 was the first training subject and was not given any tests of the transfer rake-lake series. However, since her data on the rock-lock series were reliable and similar to those of subsequent subjects, we included her in the study reported here.

2. S1 completed six consecutive blocks of 18 trials on each standard, and her performance on the first block was used as warm-up and discarded.

3. Treating the average function as if it were an individual's data, 95% confidence intervals were established for both identification performance (> 14/20 for chance probability = .50) and discrimination performance (> 12/24 for chance probability = .33) by the binomial expansion. These confidence intervals are indicated on the right-hand side of each figure.

4. $\hat{P}$(Hits) = [P(Hits) − P(False Alarms)]/[1 − P(False Alarms)]. Maximum = .50. P(Hits) = No. Correct "D"/Total D + SD trials.

5. There was no relationship between boundary location and the R standard on which subjects were trained.

APPENDIX A

Nominal parameter values for the rock-lock stimulus series are given in Table 6. The numbers represent the duration (in milliseconds) of the initial steady state (SS) and the transition (Tran) of the first formant (F1), the center frequencies of the second (F2) and third (F3) formants at the beginning of the syllables (Start), and the center frequency of F3 at the point of inflection 35 msec into the syllable (T=35).

Table 6

Stimulus Number	F1 Duration (msec)		Formant Frequencies (Hz)		
	SS	Tran	F2 Start	F3 Start	F3(T = 35)
1	14	49	1067	1477	1576
2	14	49	1083	1611	1694
3	21	42	1099	1731	1808
4	21	42	1115	1847	1915
5	28	35	1131	1972	2029
6	28	35	1147	2104	2135
7	35	28	1156	2229	2262
8	35	28	1172	2345	2362
9	42	21	1189	2466	2484
10	42	21	1207	2594	2594

Constant -ock Portion
Formant Center Frequencies (Hz)

	Vowel			Final Closure		
F1 Start	F1	F2	F3	F1	F2	F3
346	621-707	1198-1233	2557	621	1288	2104

APPENDIX B

Individual identification (ID) and oddity discrimination (Disc) functions for the eight subjects on the rock-lock stimulus series are shown in Figure 10. Open circles and dashed lines show Pretest 1 functions; open triangles and dot-dash lines represent Pretest 2 functions for S5-S8; closed circles and solid lines give posttest functions. Vertical arrows in discrimination functions indicate the identification category boundary.

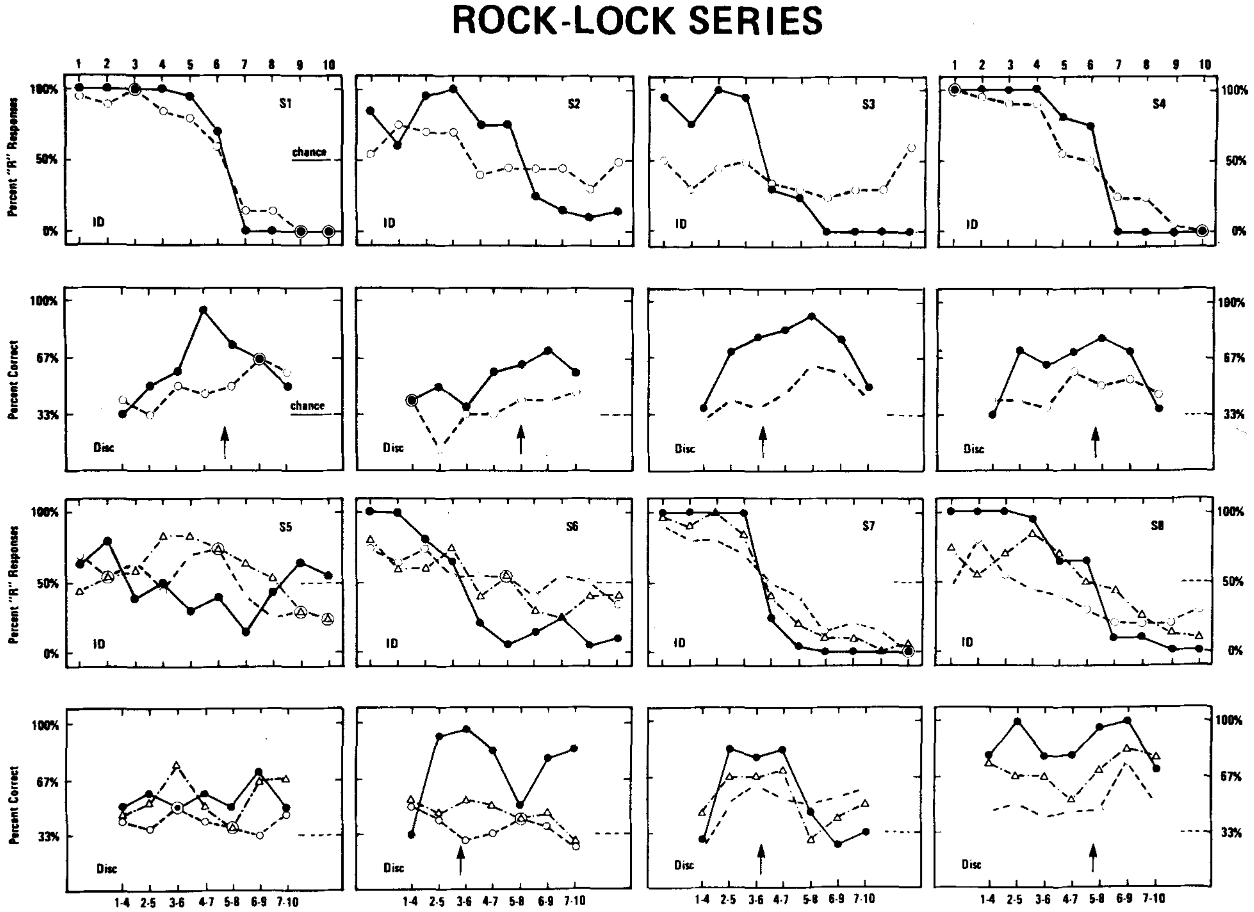


Figure 10

APPENDIX C

Nominal parameter values for the rake-lake stimulus series are given in Table 7. The numbers represent duration (in milliseconds) of the initial steady state (SS) and first formant transition (T1), the center frequencies of the first (F1), second (F2), and third (F3) formants at the beginning of the stimuli.

APPENDIX D

Individual identification (ID) and oddity discrimination (Disc) functions for the seven subjects on the rake-lake stimulus series are given in Figure 11. Open circles and dashed lines show Pretest 1 functions; open triangles and dot-dash lines represent Pretest 2 functions for S5-S8; closed circles and solid lines give posttest functions. Vertical arrows in discrimination functions indicate the identification category boundary.

Table 7

	Duration (msec)		Starting Formant Frequencies (Hz)		
Stimulus	SS	T1	F1	F2	F3
1	36	66	255	1077	1691
2	36	66	255	1153	1858
3	36	66	255	1153	2025
4	48	48	314	1153	2193
5	54	36	314	1230	2360
6	54	36	373	1230	2527
7	60	24	373	1306	2696
8	66	18	432	1306	2867
9	66	18	432	1306	3029
10	66	18	432	1382	3196

Constant -ake Portion

Starting Frequencies (Hz)			Ending Frequencies (Hz)		
F1	F2	F3	F1	F2	F3
509	2527	3029	432	2756	3029

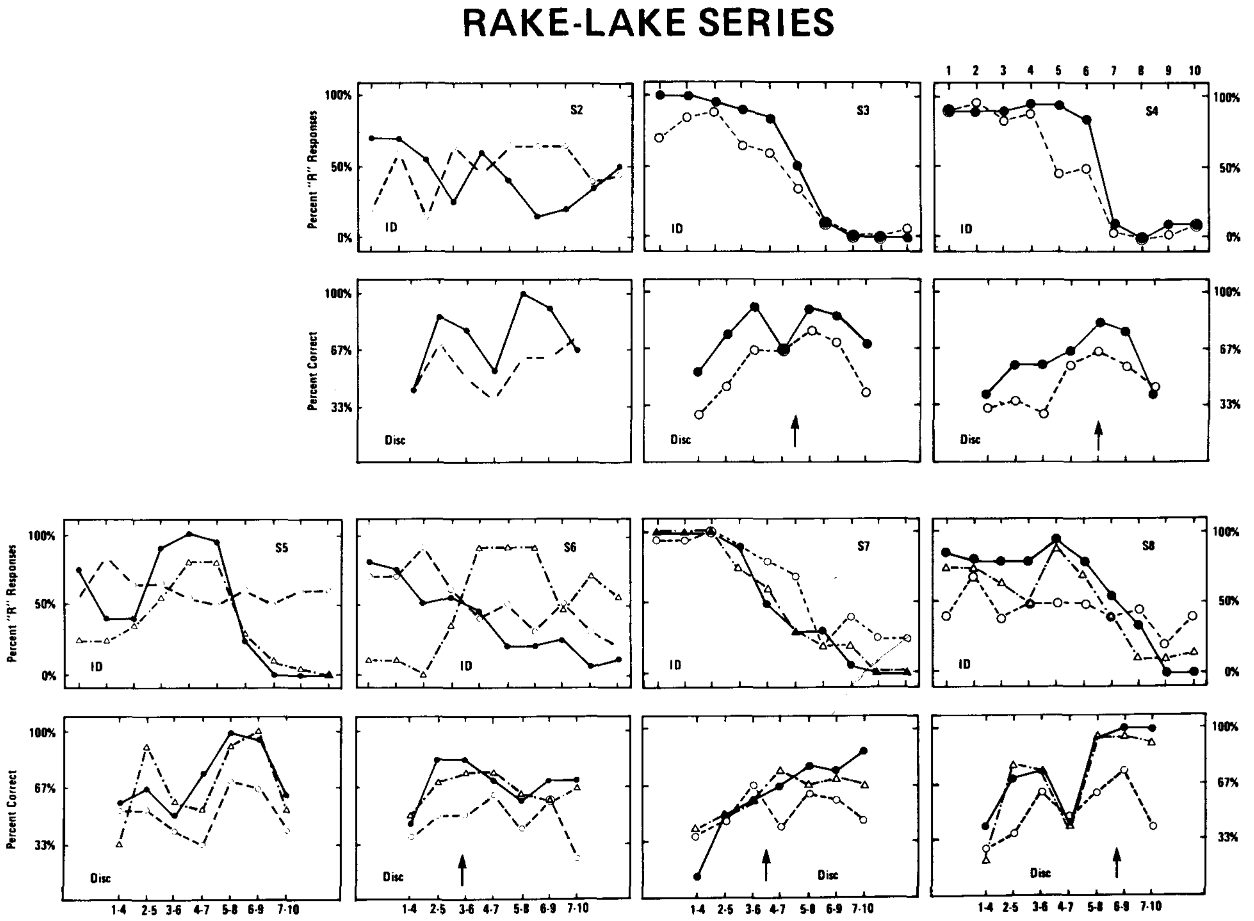


Figure 11

(Manuscript received January 26, 1984;
revision accepted for publication June 11, 1984.)