Bulletin of the Psychonomic Society
1986, 24 (6), 419-422

Training intraphonemic discrimination of /r/-/l/

WINIFRED STRANGE and LINDA POLKA
University of South Florida, Tampa, Florida

and

SIBYLLA DITTMANN
University of Minnesota, Minneapolis, Minnesota

To assess the extent to which adults can discriminate intraphonemic differences in place-of-articulation, we gave 4 subjects extensive training on a synthetic *rock–lock* series, using a same-different discrimination task. Transfer of training was evaluated using oddity discrimination tests of the training series and an acoustically dissimilar *rake–lake* series. Performance during training showed gradual improvement in intraphonemic perception, with asymptotic performance well below 100% accuracy. Oddity discrimination improved for the *rock–lock* stimuli, but not for the *rake–lake* series, suggesting that learning was specific to the acoustic parameters of the training stimuli. Distinctions between auditory, phonetic, and phonemic modes of speech perception are discussed.

Interest in the extent to which adult speakers of a language can discriminate speech sounds that fall within a single (native) phoneme category arises from two research traditions. Information-processing theorists and psychoacousticians, concerned with providing explanations of the categorical perception phenomenon, have used training studies of *intra*phonemic discrimination to argue against models of speech perception that posit a specialized linguistic mode of perception (Carney, Widin, & Viemeister, 1977; Samuel, 1977). From another tradition, researchers concerned primarily with language learning have shown that adult foreign language learners have difficulties in perception (as well as production) of nonnative phoneme contrasts that often persist even after extensive conversational instruction (Flege, 1984; Goto, 1971; MacKain, Best, & Strange, 1981). Training studies of intraphonemic discrimination using synthetic speech stimuli thus offer an experimental approach to the modification of speech perception that could have practical as well as theoretical importance.

Almost all training studies to date (Carney et al., 1977; Edman, 1981, Experiments 1 and 3; Pisoni, Aslin, Perey, & Hennessy, 1982; Samuel, 1977) have investigated the voice-onset-time (VOT) dimension. In general, these studies have shown that American English (AE) speakers learn to discriminate small (intraphonemic) differences in VOT when trained with tasks that reduce stimulus uncertainty and memory constraints. These studies also have reported nearly complete transfer of discrimination training to new synthetic stimuli that varied identically in VOT (Edman, 1981, Experiment 3), and to perceptual tasks that more closely approximate the memorial demands and stimulus uncertainty found in speech communication (Carney et al., 1977; McClasky, Pisoni, & Carroll, 1983).

Edman (1981, Experiment 2) investigated intraphonemic discrimination of acoustic dimensions underlying place-of-articulation contrasts among stop consonants. Although most subjects demonstrated improvement in intraphonemic discrimination with training, the change appeared to be more gradual than for VOT, subjects were more variable, and transfer of training to a new series was inconsistent. Edman speculated that intraphonemic discrimination was more difficult because the acoustic features differentiating place-of-articulation (differences in the rapidly changing upper spectral prominences) were less "accessible" than were VOT parameters.

In the present study, subjects were trained on intraphonemic discrimination of a synthetic series underlying the contrast between the American English liquids, /r/ and /l/. This contrast can be considered a place-of-articulation distinction, although the differences in articulatory gestures and resulting acoustic parameters are complex. Exploring possible ways to modify the perception of liquids is of practical interest because many second-language learners of English (e.g., Japanese, Korean, and others) find these speech sounds notoriously difficult to master.

In a recent study (Strange & Dittmann, 1984), native Japanese adults learning English were given extensive training with a *rock–lock* series, using an all-step same-different discrimination procedure (Carney et al., 1977). All 8 subjects showed gradual improvement, and 7 of the 8 subjects showed significant transfer of training to identification and/or oddity discrimination tests using the training stimuli. However, transfer to an acoustically dissimilar rake-lake synthetic series was inconsistent and

This research was supported by a grant, "Studies of Speech Perception" (MH-21153), to James J. Jenkins and Winifred Strange, 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. The authors wish to express their appreciation to Lenief Heimstead and Grant Miller for developing the computer hardware and software, and to James J. Jenkins for his comments on the manuscript. Requests for reprints should be sent to Winifred Strange, Department of Communicology, University of South Florida, Tampa, FL 33620.

there was little or no posttraining improvement in performance on natural speech /r/-/l/ minimal pairs.

In the present study, native English speakers were given discrimination training on the *rock–lock* series, using the same design and procedures employed with the Japanese subjects. Three questions were of interest: (1) To what extent does intraphonemic discrimination improve on the *rock–lock* series? (2) Given comparable amounts of training, is the level of final performance for native English speakers higher than that of native Japanese speakers? (3) Does training lead to improvement in performance on more demanding discrimination tests of the training stimuli, and does it transfer to tests of the *rake–lake* series?

METHOD

Subjects

Six female college students at the University of Minnesota were participants. Four were assigned to the experimental group and received training and 2 served as control subjects. All 6 subjects were monolingual, native AE speakers and reported having no hearing problems.

Stimulus Materials

Two sets of synthetic stimulus materials were used in pre- and posttraining categorical perception tests: (1) a 10-step *rock–lock* series, simulating an adult male's productions, and (2) a 10-step *rake–lake* series, simulating an adult female's productions. Each series had been utilized in previous studies and is described in detail in Strange and Dittmann (1984).

For use in the training task, the 10 stimuli of the *rock–lock* series were filtered at 4750 Hz and converted to digital waveform files (10K/sec sampling rate with 12-bit resolution), using a PDP-8E laboratory computer. Software programs controlled the reconversion of these files to analog signals that were filtered at 4750 Hz and presented via TDH-39 earphones to the subject seated in a one-person IAC acoustic chamber.

Procedures

Pre- and posttraining tests. Identification and three-step oddity discrimination tests were performed using the traditional categorical perception procedures (see Strange & Dittmann, 1984). Stimuli were presented binaurally via earphones (TDH-39) at 70 dB SPL; all subjects were first tested on the *rock–lock* series, then on the *rake–lake* series. Posttests were simply a repetition of pretest procedures, given after the 4 experimental subjects had completed training. For the 2 control subjects, this constituted a retest, completed approximately 3 weeks after initial testing.

Training. Stimuli were presented binaurally over earphones (TDH-39) at 70 db SPL in pairs with a 1-sec ISI. The first stimulus of the pair 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 once with each of the nine remaining stimuli of the series; on "same" trials, the standard was repeated. The sequence of "same" and "different" trials was randomly selected by the computer for each block. Subjects responded by pressing response buttons marked "S" and "D," and immediate feedback was given after each response.

All experimental subjects completed training with two standards: Stimulus 8 (categorized as /l/) and Stimulus 3 (for 2 subjects) or Stimulus 4 (for 2 subjects). Both Stimulus 3 and Stimulus 4 were categorized consistently as /r/. The first session for each subject, which consisted of five consecutive blocks with each of the two standards, was considered task familiarization, and results were not analyzed further. Each subject then completed 15 training sessions, taking place on separate days over the course of about 3 weeks. A training session consisted of six (for 2 subjects) or seven (for 2 subjects) consecutive blocks of 18 trials with each standard. 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 was counterbalanced across subjects.

RESULTS

All-Step Discrimination

Although subjects' initial discrimination was quite good, all 4 subjects showed some improvement over sessions on the all-step discrimination task. Averaging performance across the two standards in the two initial sessions (Sessions 1-2), the mean number of errors/block was 3.8 (out of 18), or 79% correct responses; subjects ranged from 77%-83% correct discrimination. By Sessions 14-15, the average error rate fell to 2.6 errors/block, or 86% correct responding. The false-alarm rate (errors on "same" trials) averaged 1.3 errors/block in the initial sessions and fell to 0.8 in Sessions 14-15. In general, then, training with this task resulted in an average improvement of about 7%, but performance leveled off at about 86% accuracy, ranging across the 4 subjects from 81% to 91%.

Asymptotic performance by the American group was similar to that of Japanese learners of English (Strange & Dittmann, 1984). Although performance in initial sessions for the 8 Japanese subjects averaged below that for the Americans (5.1 errors/block, or 71% accuracy), their final performance was almost the same as the Americans' (2.9 errors/block, or 84% accuracy). Individual Japanese subjects ranged from 79% to 89% correct discrimination in the final two sessions. We can conclude, then, that 15 to 18 training sessions with the all-step discrimination task were sufficient to eradicate the initial differential effect of linguistic experience on all-step discrimination of the training stimuli.

To illustrate specific improvement in intraphonemic discrimination, Figure 1 presents individual psychophysical functions. Percentages of correct "different" responses (hits) are given for each comparison stimulus paired with the R standard to the left and the L standard to the right. Percentages of false alarms are shown by the squares plotted above the standard stimulus. Performance on Sessions 1-3 and Sessions 13-15 are superimposed.

In initial training sessions (dashed functions), *inter*phonemic discrimination (i.e., comparison stimuli from across the phoneme boundary indicated by the vertical dashed line) was nearly perfect for all subjects. Except for Subject 1 on the R standard, above-chance *intra*phonemic discrimination at the beginning of training was primarily observed on comparisons involving the acoustically intermediate stimuli, 5, 6, and 7.

Improvement with further training in intraphonemic discrimination, reflected by differences between functions for initial and final sessions (solid functions), varied considerably across subjects and standards. Subject 3 showed the greatest overall improvement; Subject 2 also improved from initial to final sessions with both standards. Subject 4 showed little improvement in final performance with the R standard; Subject 1 showed no improvement on L standard comparisons. Remaining errors were primarily on intraphonemic comparison stimuli toward the endpoints of the series, that is, away from the phoneme boundary from the standard.

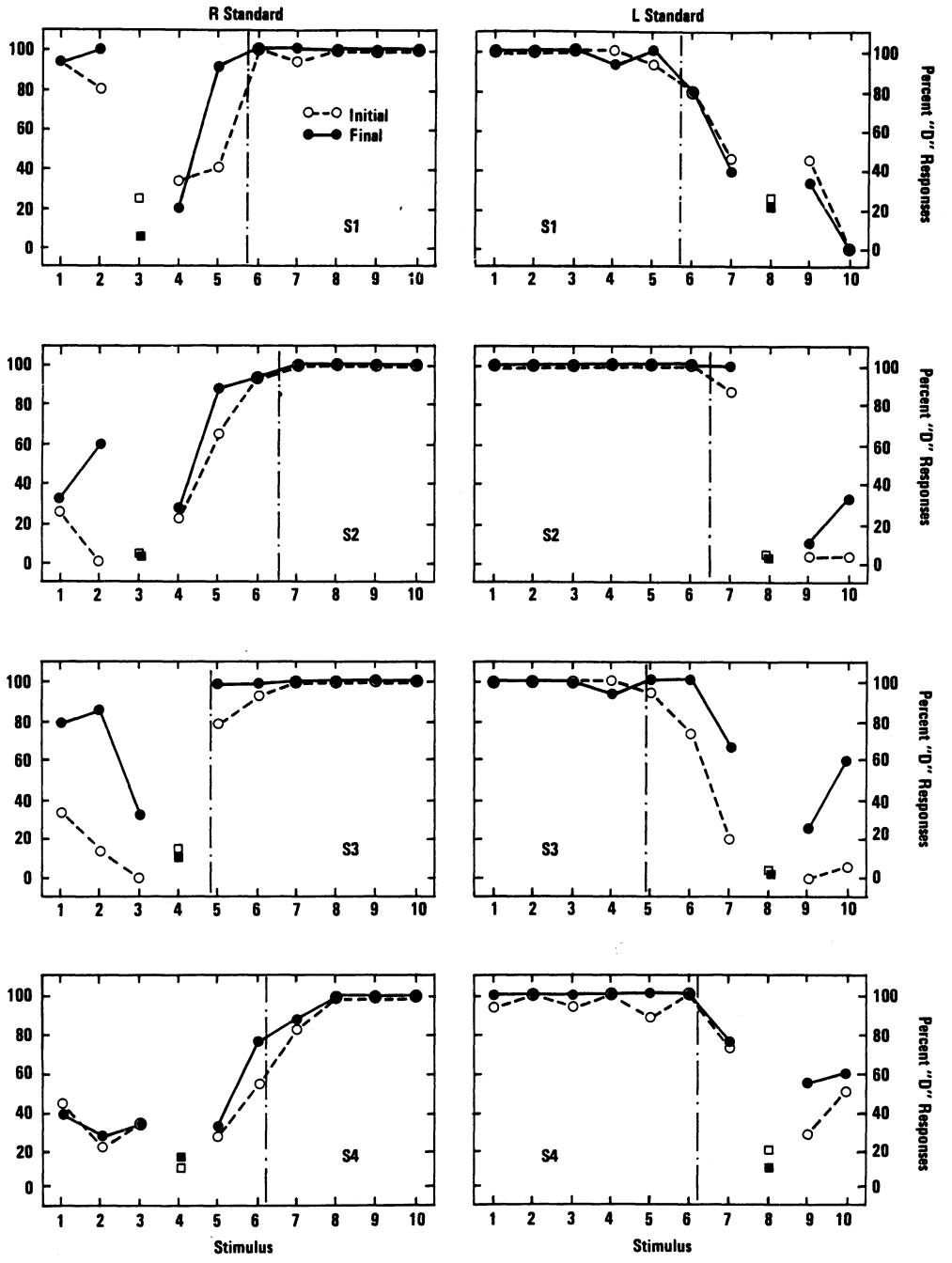

Figure 1. Psychophysical functions for individual subjects on R standard blocks (left) and L standard blocks (right) during initial three training sessions (open circles) and final three sessions (closed circles).

Identification and Oddity Discrimination

Regarding questions of transfer of training to new tasks, we looked first at pre- versus posttraining performance with the *rock–lock* stimuli. Identification boundaries did not shift appreciably as a function of training, and the steepness of the crossover between categories did not change. Training did have an appreciable effect on oddity discrimination; increases in discrimination accuracy were greatest for comparison pairs that represented cross-category comparisons for subjects. In addition, there was improvement by all 4 subjects in discrimination of Stimuli 1–4. Analyses of each individual's overall performance on pre- versus posttests showed significant improvement in discrimination accuracy for Subject 2 ($\chi^2 = 9.41$, $p < .01$), Subject 3 ($\chi^2 = 5.44, p < .05$), and Subject 4 ($\chi^2 = 6.64, p < .01$). This improvement in oddity discrimination on the *rock–lock* series can be attributed directly to the subjects' training experience.

With regard to transfer of training to oddity discrimination of the *rake–lake* series, 3 of the 4 subjects showed

no improvement from pretest to posttest. Only Subject 3 showed a significant increase in discrimination accuracy ($\chi^2 = 7.31, p < .01$). Posttraining discrimination levels for the experimental subjects as a group (66%) were somewhat better on the average than control subjects' (62%). However, they were not better than control subjects' initial discrimination test results (68% correct), which were higher than their own retest performance.

DISCUSSION

The results of this study indicate that adult English speakers are capable of discriminating, in a complex acoustic series, intraphonemic differences that underlie the distinction between AE prevocalic /r/ and /l/. With 15 training sessions, involving over 3 thousand trials, intraphonemic discrimination improved, especially for comparison pairs adjacent to the subjects' identification boundaries. However, asymptotic levels of performance were well below 100%. Improvement in discrimination of the *rock–lock* stimuli transferred to the more difficult oddity discrimination paradigm for 3 of the 4 subjects. However, there was little generalization to the *rake–lake* series, which varied the same phonetic parameters but was acoustically dissimilar to the training series.

These results support the contention of Edman (1981) that the acoustic-phonetic parameters differentiating place-of-articulation contrasts are somehow "less accessible" to subjects who are trying to make nonphonemic discriminations. This conclusion is also supported by the work of Werker and Tees (1983, 1984) who found that, although limited training (25 trials) enabled English-speaking subjects to discriminate a Hindi voicing distinction (voiceless aspirated vs. breathy voice /t^h/–/d^h/), more extensive laboratory training (500 trials) and/or 1 year of Hindi language training did not lead to accurate discrimination of a Hindi place-of-articulation distinction (retroflex vs. dental /ṭ/–/t̪/). Thus, we must be careful in generalizing results of training studies with particular acoustic-phonetic contrasts to the overall question of the ease with which phonetic perception may be modified in the laboratory or the language classroom.

The present data illustrate that the type of task employed in testing discrimination influences to a large extent the level of intraphonemic discrimination. This has been interpreted as reflecting different task-induced levels or modes of perceptual processing of speech stimuli (Fujisaki & Kawashima, 1969; Pisoni, 1973). Werker and Tees (1984) have most recently distinguished three modes or factors in perceptual processing: a phonemic or phonological mode (previously called phonetic), in which subjects respond in accordance with language-specific (learned) categories; an auditory mode in which performance is based on briefly available sensory information, and a third mode, which they label phonetic, wherein discrimination is based on "natural phonetic boundaries" (i.e., phonetically relevant acoustic parameters that are functional (phonemic) in some languages, but not in the subjects' native language.)

In our study, performance on the all-step task even at the outset reflected subjects' ability to make intraphonemic discriminations. Thus, we can hypothesize that they were operating in either a phonetic or auditory mode. By the end of training, American and Japanese subjects (Strange & Dittmann, 1984) were performing at about the same level. Thus, the effect of phonological experience was no longer influencing performance on this task. However, for both groups, discrimination of intraphonemic comparisons adjacent to the phonetic boundary (i.e., the acoustically intermediate stimuli) tended to be better than for intraphonemic comparisons on the opposite side of the standard from the boundary (i.e., the acoustic extremes of the series). In other words, there appeared to be a "phonetic boundary effect" for both sets of subjects, only one of which had extensive experience with the contrast as a phonological distinction.

We suggest that the American (and perhaps even the Japanese) subjects may have been performing on the basis of discrimination of three "phonetic" categories: clear case /r/, clear case /l/, and an intermediate phonetic category. With acoustic synthesis, it is not possible to relate the intermediate variants of the stimulus series to articulatory gestures involved in liquid production. However, the intermediate stimuli can be heard as a nonlabialized /w/ or a distorted /r/. Acoustic analysis of approximant consonants of other languages might shed light on whether there are "natural" phonetic boundaries along this acoustic continuum other than the one used in English.

Transfer to the oddity test, with its increased memory demands and stimulus uncertainty, argues for an interpretation that subjects were operating at the phonetic level. However, the overall lack of transfer to the *rake–lake* series argues against such an interpretation. Further studies designed specifically to differentiate these modes of perception must be accomplished before a definitive answer can be given (see Werker & Logan, 1985).

We can conclude on the basis of this study that adults are capable of making intraphonemic discriminations and that training with feedback enhances performance on the particular stimuli being used. Furthermore, we can conclude, in agreement with other researchers (Aslin, Pisoni, & Jusczyk, 1983; Werker & Tees, 1984) that differential linguistic experience does *not* lead to any permanent sensory-neural change in the ability to discriminate phonetically relevant acoustic parameters. Rather, linguistic experience shapes subjects' access to those discrimination abilities in distinguishing phonetic categories. The extent to which a person, having learned a particular phonological system, can learn to make functional use of nonnative phonetic contrasts remains a complex question. It appears that all types of contrasts are not equivalent in terms of the ease with which modification in adulthood can occur.

REFERENCES

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

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. (1981). Learning of intraphonemic discrimination for several synthetic speech continua. *Dissertation Abstracts International*, **41**, 4291B. (University Microfilms No. 81-09, 419)

Flege, J. E. (1984). The effect of linguistic experience on Arabs' perception of the English /s/ vs. /z/ contrast. *Folia Linguistica*, **18**, 117-138.

Fujisaki, H., & Kawashima, T. (1969). Some experiments on speech perception and a model for the perceptual mechanism. *Annual report of the Engineering Research Institute* (Vol. 28). Tokyo: University of Tokyo, Faculty of Engineering.

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

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

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

Pisoni, D. B. (1973). Auditory and phonetic codes in the discrimination of consonants and vowels. *Perception & Psychophysics*, **13**, 253-260.

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 & Performance*, **8**, 297-314.

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

Strange, W., & Dittmann, S. (1984). Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Perception & Psychophysics*, **36**, 131-145.

Werker, J. F., & Logan, J. S. (1985). Cross-language evidence for three factors in speech perception. *Perception & Psychophysics*, **37**, 35-44.

Werker, J. F., & Tees, R. C. (1983). Developmental changes across childhood in the perception of nonnative speech sounds. *Canadian Journal of Psychology*, **37**, 278-286.

Werker, J. F., & Tees, R. C. (1984). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America*, **75**, 1866-1878.

(Manuscript received for publication June 20, 1986.)