Applied Psycholinguistics (1981) 2, 369–390
Printed in the United States of America

Categorical perception of English /r/ and /l/ by Japanese bilinguals

KRISTINE S. MacKAIN
Haskins Laboratories
and Cornell University Medical College

CATHERINE T. BEST
Haskins Laboratories
and Teachers College, Columbia University

WINIFRED STRANGE
University of Minnesota

ADDRESS FOR CORRESPONDENCE
Dr. Kristine S. MacKain, Department of Psychiatry, Cornell University Medical College, 525 East 68th Street, New York, New York 10021.

ABSTRACT
Categorical perception of a synthetic /r/–/l/ continuum was investigated with Japanese bilinguals at two levels of English language experience. The inexperienced Japanese group, referred to as Not-experienced, had had little or no previous training in English conversation. The Experienced Japanese group had had intensive training in English conversation by native American-English speakers. The tasks used were absolute identification, AXB discrimination, and oddity discrimination. Results showed classic categorical perception by an American-English control group. The Not-experienced Japanese showed near-chance performance on all tasks, with performance no better for stimuli that straddled the /r/–/l/ boundary than for stimuli that fell in either category. The Experienced Japanese group, however, perceived /r/ and /l/ categorically. Their identification performance did not differ from the American-English controls, but their overall performance levels on the discrimination tests were somewhat lower than for the Americans. We conclude that native Japanese adults learning English as a second language are capable of categorical perception of /r/ and /l/. Implications for perceptual training of phonemic contrasts are discussed.

Languages differ in their phonological and phonetic inventories. For example, in a particular language (L1), two phones may occur, while in another language (L2), the phones may not appear at all. Or, L1 and L2 may share two phones, but in L1 the phones may be phonologically contrastive, while in L2, they may occur in contextual or free variation rather than being used to distinguish meaning. Because of this variation across languages, several questions have been asked about the potential role of linguistic experience in the perception of phonological categories. Are speakers universally sensitive to the parameters that distinguish phonological contrasts in all languages, or does experience with the phonological categories of one's native language

 0145-7164/81/040369-22/$2.50

affect the perception of those contrasts? For native speakers of languages that do not make use of particular speech sounds as a phonological contrast, is the perception of those sounds affected? If so, can perception of a phonetic contrast be modified in adulthood through learning a language that does employ the contrast as a phonological opposition?

The first two questions have been answered to some extent by cross-language investigations of vowel and consonant perception. It has been found that linguistic experience with phonological contrasts does affect perception of them, at least for some vowels and consonants. For vowels, experience influences perceptual discrimination judgments made along an interval scale, but does not produce differential nominal judgments. Using nominal (same/different) judgments for pairs of stimuli, Stevens, Liberman, Studdert-Kennedy, and Öhman (1969) found no difference in the ability of native Swedish and American-English speakers to detect differences in vowel contrasts that were phonologically distinct in Swedish but not in English. However, by employing a more sensitive interval scale discrimination measure, in which listeners judged the relative magnitudes of differences they perceived between members of various pairs of vowel stimuli, Terbeek (1977) found that language experience in monolinguals of five different languages does affect vowel perception. The perceptual distance between two vowels was judged to be much greater if the pair contrasted phonologically in the subjects' native language than if the pair was not a native contrast.

Linguistic experience also affects the location of phonetic perceptual boundaries between stop consonant contrasts. For instance, Voice Onset Time (VOT) – the time between release of articulatory closure and onset of phonation – is a sufficient cue for phonological categorization of stop consonants in perception (Lisker & Abramson, 1970) and production (Lisker & Abramson, 1967). These investigators found cross-language differences in the location of the perceptual boundary between "voiced" and "voiceless" phonetic categories along a synthetic stimulus continuum underlying VOT. For each language group, identification (Lisker & Abramson, 1970) and discrimination (Abramson & Lisker, 1970) responses were generally in close correspondence. Moreover, identification and discrimination responses for Thai and American-English speakers were different and generally corresponded to their respective stop voicing production distributions, reported in an earlier study (Lisker & Abramson, 1964). Similar effects of experience have been found with native Spanish speakers (Abramson & Lisker, 1973; Williams, 1977) whose VOT production distributions differ from both Thai and English. It appears, then, that experience with specific voicing contrasts among stop consonants determines the location of perceptual boundaries separating those phonological categories along the articulatory continuum.

The effects of linguistic experience just summarized suggest, in addition, the converse situation – that lack of experience with a given phonological contrast should result in a poorly-defined perceptual boundary separating the two members of that contrast. Cross-language studies on categorical perception of nonnative phonetic contrasts have addressed this issue. Categorical perception is said to occur if the subject cannot discriminate speech sounds any better than she/he can identify them within different phonological categories. Under these conditions, equal increments along a phonetically relevant articulatory continuum (with acoustic consequences) are not discriminated unless the increment crosses the boundary between phonetic categories (e.g., Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967).

In this vein, recent studies (Miyawaki, Strange, Verbrugge, Liberman, Jenkins, & Fujimura, 1975; Mochizuki, 1981) have assessed the perception of synthetic /r/–/l/ continua by native Japanese and native American-English speakers. Native Japanese speakers who have learned English as a second language in adulthood are notorious for having difficulty in discriminating /r/ from /l/. In spoken Japanese, the liquid /l/ does not occur. Although a form of /r/ ("rhotic") is said to occur phonemically, it fits the criteria of a flap [ɾ], and is more similar acoustically and articulatorily to the American-English voiced dental-alveolar flap [ɾ] than to the approximant [ɹ] in American English (Miyawaki, 1973; Price, 1981). The 21 Japanese subjects in the Miyawaki et al. (1975) experiment had all studied English for at least 10 years; however, their instruction did not stress conversational English and only two subjects had resided in an English-speaking country. The Japanese subjects completed an oddity discrimination task on a synthetic /r/–/l/ continuum that varied only the pattern of time-varying frequency change in the third oral formant (F3), which is seen in a speech sound spectrogram as the band of resonance energy that is third from the lowest-frequency band running along the time axis. Differences in the time-varying pattern of F3 are considered to be the primary cue for the /r/–/l/ contrast in English. For syllable-initial /r/, F3 begins at a frequency lower than its steady-state vowel frequency and rises during the initial consonantal portion of the syllable, whereas F3 for syllable-initial /l/ usually begins at a higher-than-steady-state frequency and falls slightly during the initial portion of the syllable. Presumably, neither endpoint of the Miyawaki et al. continuum corresponded to the spectral configuration of the Japanese /r/ category. American-English subjects completed both oddity discrimination and forced-choice identification tasks. The former task required a decision as to which of three stimuli per trial was different from the other two stimuli, which were identical. The Americans showed typical categorical perception results: they divided the continuum consistently into two phonetic categories in the identification task, and discriminated between-category stimulus comparisons well but within-category comparisons poorly. In contrast, the Japanese did not discriminate the series categorically: discrimination was nearly random and was no better for comparisons that crossed the phonetic boundary than for those lying within either the /r/ or the /l/ category.

Whereas the Miyawaki et al. (1975) study included a test of /r/–/l/ discrimination by Japanese, Mochizuki (1981) tested nine Japanese speakers only on an identification test, which used a synthetic /r/–/l/ series (again, only the F3 spectral configuration was varied). Her Japanese subjects divided the continuum into two distinct phonetic categories with a perceptual boundary that closely corresponded to that of an American-English control group; this would seem at odds with the Miyawaki et al. report. However, it may be important that the English language experience of the Japanese in the two experiments differed somewhat. Although both sets of subjects had had similar levels of formal training with English, Mochizuki's subjects had all lived in an English-speaking country, and were still residing there at the time of testing (range = 6 months to 4 years in U.S.).

The present investigation examined categorical perception of /r/ and /l/ by native Japanese speakers at two levels of English language experience, and compared their performance to that of native American-English speakers. The design was a replication and extension of Miyawaki et al. (1975) with the following changes: (1) The synthetic stimulus series included variation in both spectral and temporal acoustic dimensions that differentiate natural American-English /r/ and /l/ (Dalston, 1975). These redundantly cued stimuli were used in order to optimize the Japanese subjects' opportunity to show perceptual differentiation of the /r/–/l/ contrast. (2) In addition to the oddity discrimination task used in previous studies, an AXB discrimination task was included. This task requires the subject to judge whether the middle stimulus (X) on each three-item trial is identical to the first (A) or third (B) stimulus; A and B vary from trial to trial, but on each trial they always differ acoustically by a fixed amount along the stimulus continuum. The AXB task has lower memory demands than the oddity task (defined earlier), and is thought to provide a better opportunity for detecting auditory differences. (3) An absolute identification task was included for computing predicted discrimination performance by the American and the two Japanese groups. These three tasks provide an extensive perceptual profile for the phonological contrast with stimuli that closely resemble natural speech exemplars of the phonological categories in American English. The primary question of interest was whether Japanese-English bilinguals with relatively intensive experience conversing with native English speakers would identify and discriminate /r/ and /l/ according to American-English categories, while Japanese with less English conversation experience would show less categorical /r/–/l/ perception.

METHOD

Subjects. The American group was comprised of 10 undergraduates (5 males, 5 females) recruited through notices posted on campus bulletin boards at Yale University. We recruited Japanese adults from the Yale community by telephone; 12 agreed to participate (7 males and 5 females). All were Japanese natives who had moved to the U.S. as adults, except for one young woman who had moved at age 15. They filled out a language-experience questionnaire prior to the experiment, and two subgroups were chosen on the

Table 1. *American English conversation experience of the Experienced and the Not-experienced Japanese subjects*

	Factors *Percent of day conversing in English since coming to U.S. (25, 50, 75, or 100%)*	*Number of hours per week in instruction on English conversation by native speakers*	*Number of months experience in English conversation with native speakers*
Experienced Japanese			
S 7 (♀)a	75	8	48
S 8 (♂)b	75	10	48
S 9 (♂)a	75	8	18
S10 (♀)b	25	10	18
S11 (♀)c	25	4	6
	$\bar{X}$ = 55	$\bar{X}$ = 8	$\bar{X}$ = 27.6
Not-experienced Japanese			
S 1 (♀)c	25	3	5
S 2 (♂)d	25	0	5
S 3 (♂)d	25	3	2
S 4 (♂)d	25	0	6
S 5 (♀)c	25	0	18
S 6 (♂)d	25	0	18
S12 (♂)d	50	0	6
	$\bar{X}$ = 28.6	$\bar{X}$ = .86	$\bar{X}$ = 8.7

a = graduate student; b = undergraduate student; c = homemaker; d = postdoctoral associate.

basis of the amount and quality of their English conversation experience (see Table 1). The Experienced group contained 5 subjects (2 males, 3 females) who had had intensive English conversation training by native American-English speakers. The other 7 (5 males, 2 females) were designated Not-experienced, by contrast, because they had had little or no native English conversation training. All subjects reported normal hearing in both ears. Pay for participation was $3.25 per hour.

Stimuli. A /rɑk/–/lɑk/ ("rock"–"lock") series was generated on the OVE-IIIc synthesizer at Haskins Laboratories. The endpoint stimuli were traced from spectrograms of /rɑk/ and /lɑk/ utterances by an American male. Although the F3 initial steady-state and transition direction is a sufficient minimal cue for the perception of the initial /r/–/l/ contrast by Americans (O'Connor, Gerstman, Liberman, Delattre, & Cooper, 1957; Miyawaki et al., 1975), the stimulus series used here included variations not only in spectral characteristics of F3, but also in spectral characteristics of F2 and in temporal characteristics of F1. Figure 1 provides a schematic spectrographic representation of the stimuli. The series contained 10 nearly equal steps (not exactly equal because of the hardware limitations of the OVE-III synthesizer; but in all cases, the deviations from exact equality in step sizes

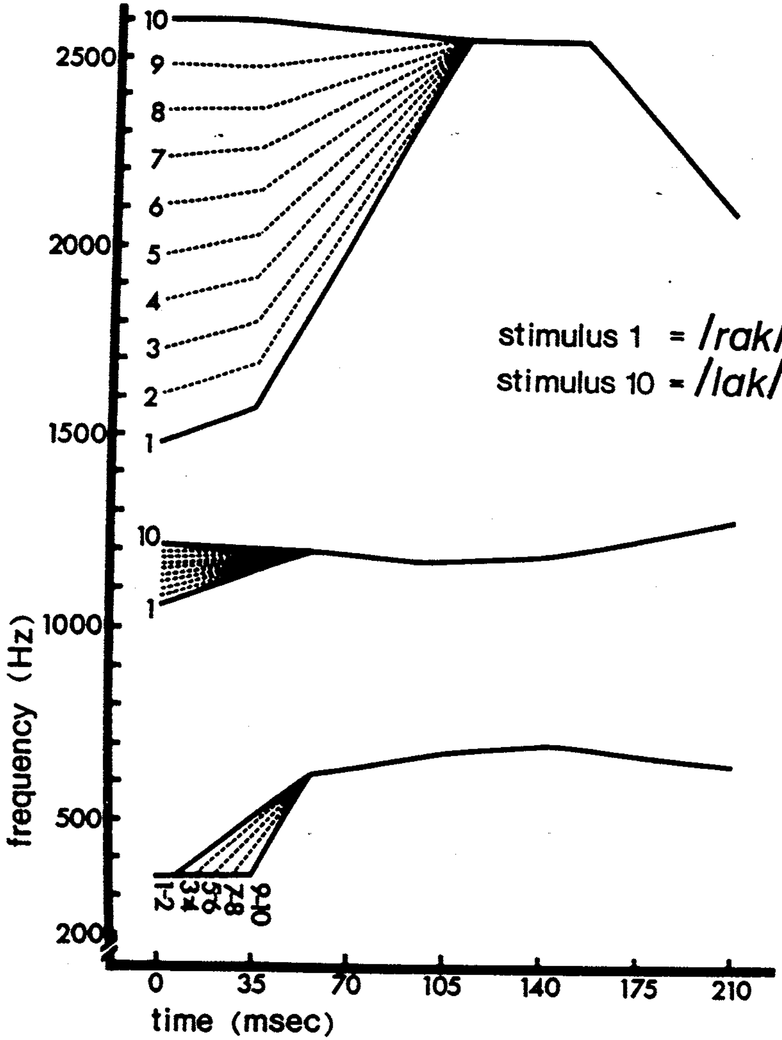

Figure 1. Schematic spectrogram representations of the 10 stimuli in the synthetic /rak/–/lak/ series.

were only a few Hz) of concurrent change for the F3 onset frequency (between 1477 Hz and 2594 Hz for /r/ and /l/, respectively), and for F3 frequency at the point of inflection (between 1067 Hz and 1207 Hz). There were 5 equal steps of F1 transition abruptness (between 21 ms and 49 ms) so that each F1 configuration occurred in 2 of the stimuli in the series. (See Appendix A for a detailed specification of stimulus parameters.)

Procedure. All subjects took part in 3 tests during a single session: (1) forced-choice identification; (2) AXB discrimination; and (3) oddity discrimination. Testing was conducted in a sound-attenuated chamber, with stimuli presented at a comfortable listening level (approximately 75 dB SPL) over TDH-39 headsets to groups of two to four subjects. The identification test consisted of 20 repetitions of the 10 stimuli, presented singly and randomized within each block of 10 trials. Intertrial intervals (ITIs) were 2.5 seconds, and interblock intervals (IBIs) were 4 sec. Subjects were asked to write "R" or "L" on each trial to indicate whether they had heard "rock" or "lock," respectively, and to choose the closer word for any ambiguous-sounding stimulus. These and subsequent instructions were typed in English for the Japanese subjects to read.

Subjects then completed an AXB discrimination test that contained 10 repetitions of each of the 2 AXB orders for the 7 possible pairings of stimuli that differed by 3 steps along the continuum (i.e., AXB pairings for the following stimulus numbers: 1–4, 2–5, 3–6, 4–7, 5–8, 6–9, 7–10). Trials were blocked by 14 (2 orders × 7 AXB pairings), and were randomized within blocks. Within-trial interstimulus intervals (ISIs) were 1 sec, ITIs were 3 sec, and IBIs 6 sec. Subjects indicated for each trial whether the second item (X) matched the first (A) or third item (B).

Next, the subjects completed the oddity discrimination test, which contained 8 blocks of 21 trials randomized across blocks of 2. Each set of 2 blocks contained one each of the 6 oddity orders for the 7 possible 3-step pairings listed above for the AXB test. There were thus 24 trials for each of the comparison pairs. On each trial, 2 of the stimuli were identical and the other differed from them by 3 steps along the continuum, as in the AXB test; however, in the oddity test trials, the "odd" stimulus member could fall in either the first, second, or third position on any 3-item trial. The subjects indicated whether the odd stimulus on each trial was the first, second, or third stimulus in the trial.

RESULTS

Americans. The Americans showed classic categorical perception of /r/ and /l/ (Figure 2). Their identification responses (left-hand panel) showed a sharp category boundary near stimulus 5, and the endpoint stimuli (1 and 10) were identified with perfect consistency as /rɑk/ and /lɑk/, respectively.

Predicted discrimination functions were computed from the identification data, for both the AXB and oddity tests. For each discrimination test, distinct peaks in performance were obtained near the /r/–/l/ category boundary (center and right-hand panels, Figure 2). The data were analyzed by a two-way Stimulus Pairs (7 levels) × Functions (2 levels: obtained and predicted) analysis of variance (ANOVA). The Stimulus Pairs effect indicated that between-category performance peaks were higher than within-category performance on both the AXB test, $F(6,54) = 11.45$, $p < .001$, and

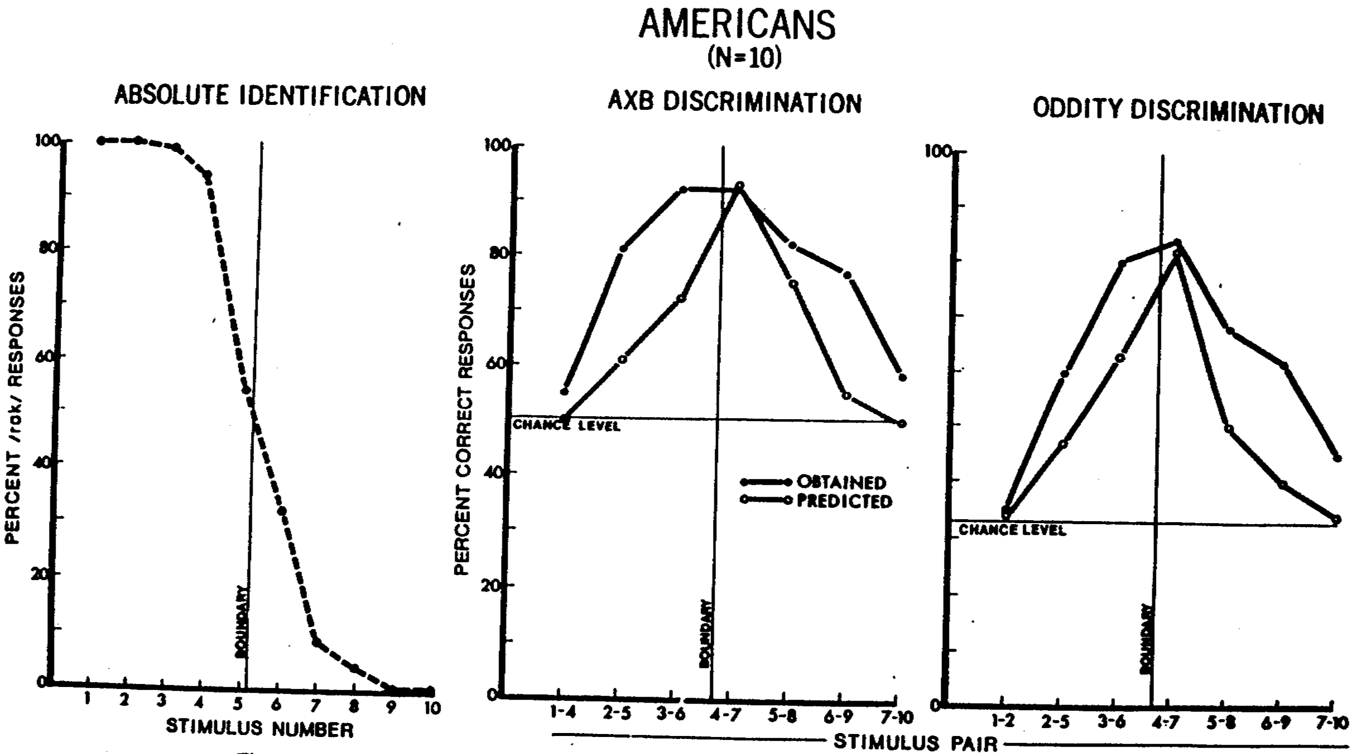

Figure 2. Response functions for the American group on the identification, AXB, and oddity tests.

the oddity test, $F(6,54) = 9.50$, $p < .001$. Obtained performance was somewhat better than predicted (solid vs. dotted lines, Figure 2), according to the Functions effect for both the AXB test, $F(1,9) = 8.44$, $p < .025$, and the oddity test, $F(1,9) = 5.25$, $p < .05$. However, simple effects tests of the Function effect for each discrimination pair (Glass & Stanley, 1970) revealed significant obtained-predicted differences only for AXB comparisons of clear-case stimuli against ambiguously identified boundary stimuli (i.e., AXB pairs: 2–5, $F(1,63) = 3.84$, $.05 < p < .10$; and 3–6, $F(1,63) = 4.83$, $p < .05$; and 6–9, $F(1,63) = 5.99$, $p < .025$; all other AXB pairs and all oddity pairs showed ns obtained-predicted differences). Thus, obtained AXB performance was no better than predicted for clear between-category comparisons (pair 4–7) and for clear within-category comparisons (1–4, 7–10). That is, obtained discrimination performance significantly exceeded category-based predictions only for AXB comparisons in which there were "category goodness" differences between stimuli within a phonetic category.

Not-experienced Japanese. In striking contrast to the Americans, the identification data indicate poor /r/–/l/ classification by the Japanese with little English conversational experience (left-hand panel, Figure 3). Category judgments hovered near chance (50%) throughout the stimulus series, and even the endpoint stimuli were, on the average, only slightly differentiated perceptually (60% vs. 40% /rak/ responses).

As predicted by their identification results, the Not-experienced Japanese performed little better than chance on the two discrimination tests. Although obtained performance on both tests (center and right-hand panels, Figure 3) appears to be slightly better than predicted, the Stimulus Pairs × Functions ANOVA on these data failed to show any significant differences. Thus, the data from this group replicate and extend the Miyawaki et al. (1975) results.

Experienced Japanese. While the results for the Not-experienced group support the Miyawaki et al. (1975) suggestion that lack of experience with /r/–/l/ as a native phonological contrast limits the perception of that contrast, the data nonetheless pose some questions: Can the limitation in /r/–/l/ perception be overcome by adults, and if so, to what extent, and through what possible types of experience? The data for the Experienced Japanese (Figure 4) address these questions. All of these subjects had had intensive English conversation training with native American-English speakers and as a group spent a larger percentage of their average day conversing in English than did the Not-experienced Japanese (see Table 1). As shown in Figure 4, their identification data (left-hand panel, Figure 4) are quite similar to the American results, and contrast with those for the Not-experienced Japanese.

In addition, the discrimination functions for these subjects, on both the AXB and oddity tests, were more similar to those of the Americans than

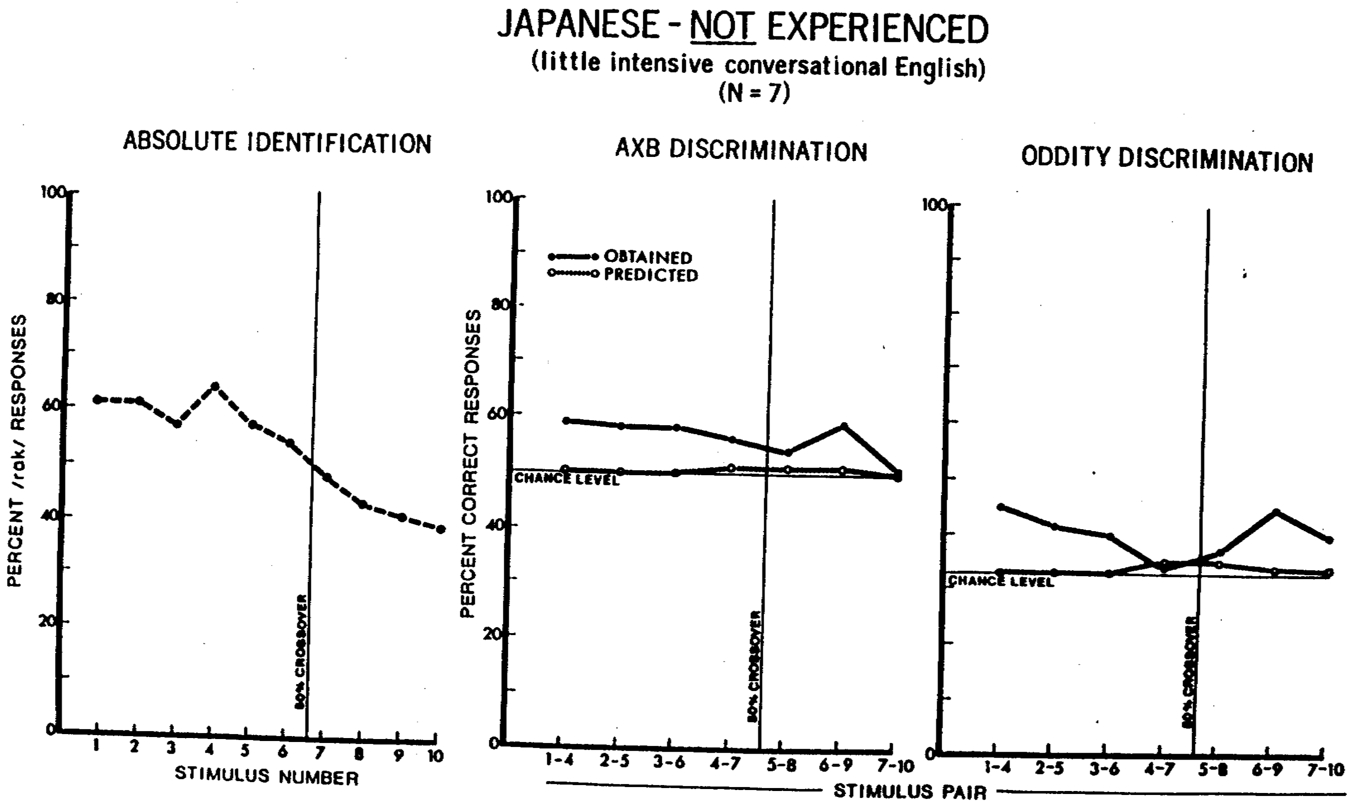

Figure 3. Response functions for the Not-experienced Japanese group on the identification, AXB, and oddity tests.

those of the Not-experienced Japanese (see group comparisons, Figure 5). Although their discrimination performance was not as high as that of the Americans, both discrimination tests revealed an increase in correct performance near the /r/–/l/ category boundary. The Stimulus Pair effect for the ANOVA on this group's discrimination data confirmed the significance of this discrimination peak on both the AXB test, $F(6,24) = 3.98$, $p < .01$, and the oddity test, $F(6,24) = 6.92$, $p < .001$. Unlike the Americans, however, obtained discrimination was not significantly better than predicted. The Stimulus Pairs × Function interaction for their AXB data, $F(6,24) = 2.70$, $p < .05$, indicated that the between-category obtained function was significantly flatter than predicted (i.e., less distinct peak).

The contrasts and similarities in the identification functions for the three groups (shown in Figure 5) suggest that the occurrence and abruptness of an /r/–/l/ category boundary for the Experienced Japanese might be related to their greater conversational English experience, relative to the other Japanese group. To assess this possibility, a measure of the steepness of the category boundary was devised, for correlation with the English language experience factors listed in Table 1. For all individuals in each group, a narrow-range PROBIT analysis of the identification data was used to fit the best ogive to the 50% crossover region of the /rɑk/–/lɑk/ categorization function (see Figure 6). These analyses included the stimulus number closest to the individual's crossover, plus the adjacent higher- and lower-numbered stimuli. The ogives fit the data well – the χ^2 values failed to approach the 5.0 value (χ^2 range = 0.0–3.8) that would denote significant deviation between obtained data and fitted ogive at the .05 alpha level, with one minor exception for the least experienced subject in the Experienced group (S11: $\chi^2 = 5.21$).

Reciprocals of the slopes of these ogives were determined as a reflection of the abruptness and direction of the perceptual category change. These values range from a theoretical minimum of 0.0, a perfectly vertical shift from 100% to 0% /rɑk/ identifications, to a maximum of 1.0, an equally abrupt but phonetically inappropriate shift from 0% to 100% /rɑk/ responses. Very small slope values thus reflect a sharp and phonetically appropriate category boundary, whereas values at .5 represent a flat slope (no true boundary), and values greater than .5 would represent a phonetically incorrect category shift.

The boundary slopes for the five Experienced Japanese were nearly as small ($M = 0.038$; range = 0.016 to 0.082) as for the Americans ($M = 0.016$; range = 0.01–0.025), while those for the seven Not-experienced Japanese were noticeably larger ($M = 0.346$; range = 0.098 to 0.708). If there were a significant positive effect of English conversation experience upon the development of clear /r/–/l/ phonetic categories by the Japanese subjects, a strong negative correlation should be found between the boundary slope and the amount of experience. All three English experience factors listed in Table 1 showed a moderate to substantial negative correlation with boundary slopes, but Factor B (number of hours per week of English

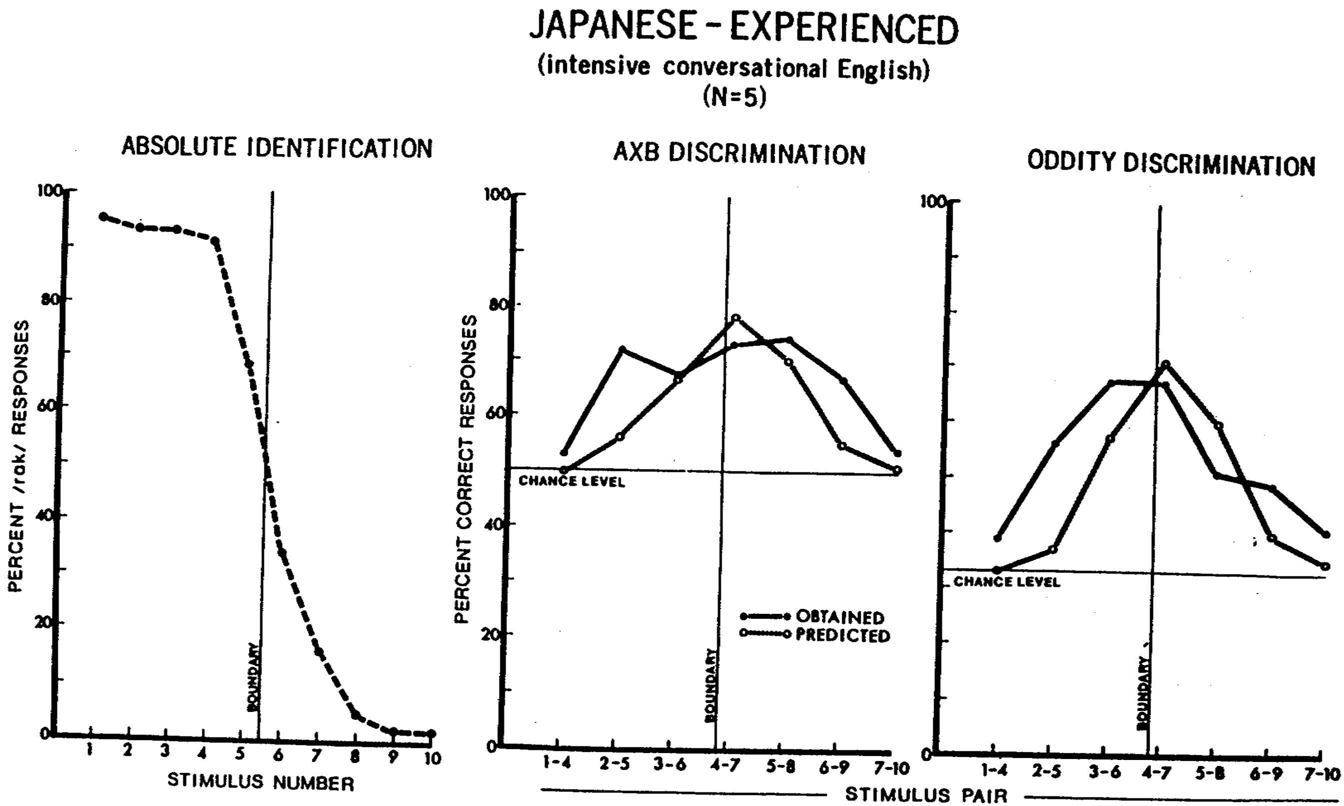

Figure 4. Response functions for the Experienced Japanese group on the identification, AXB, and oddity tests.

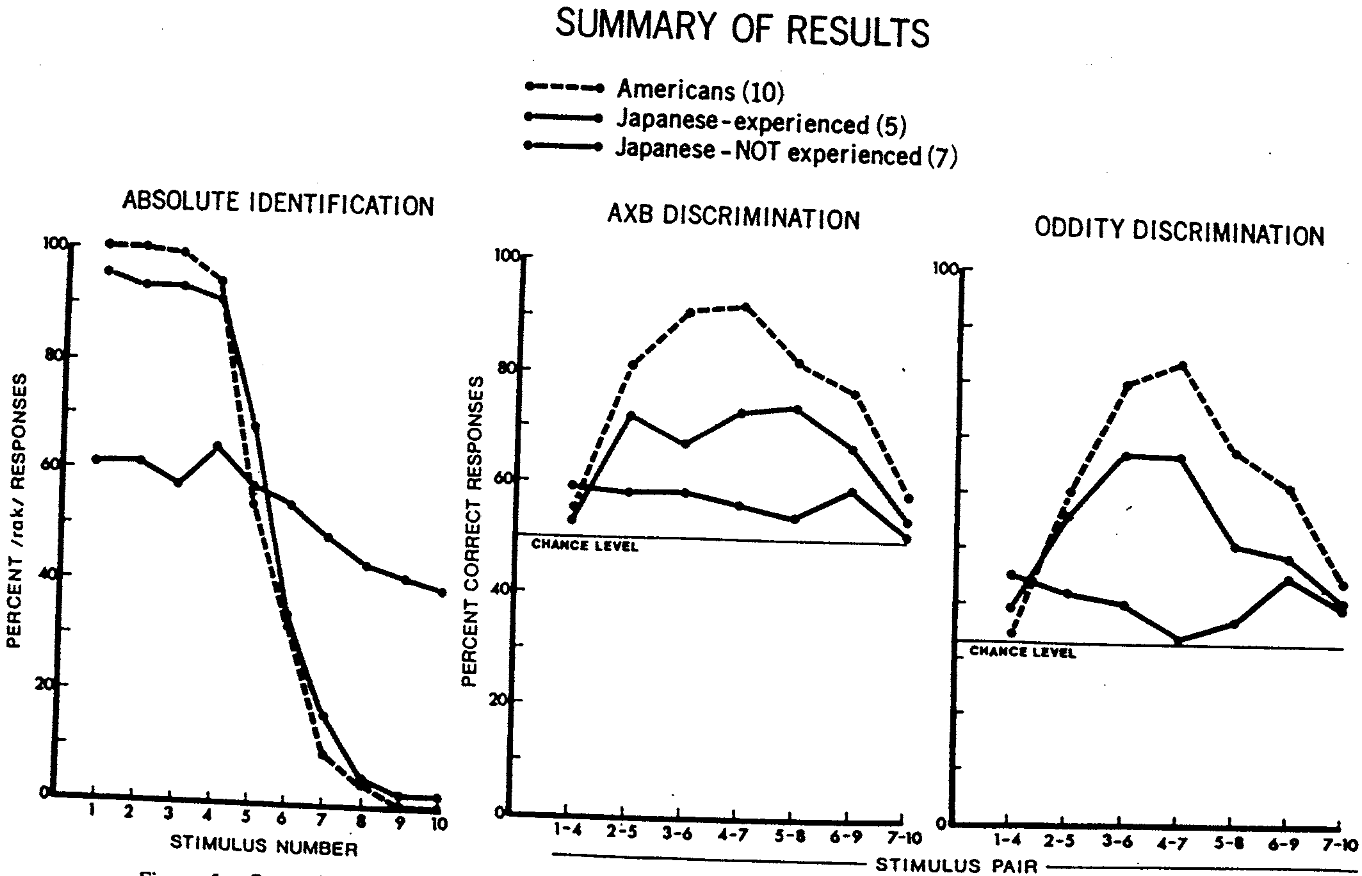

Figure 5. Comparison of Americans, Experienced Japanese, and Not-experienced Japanese results on the three tests.

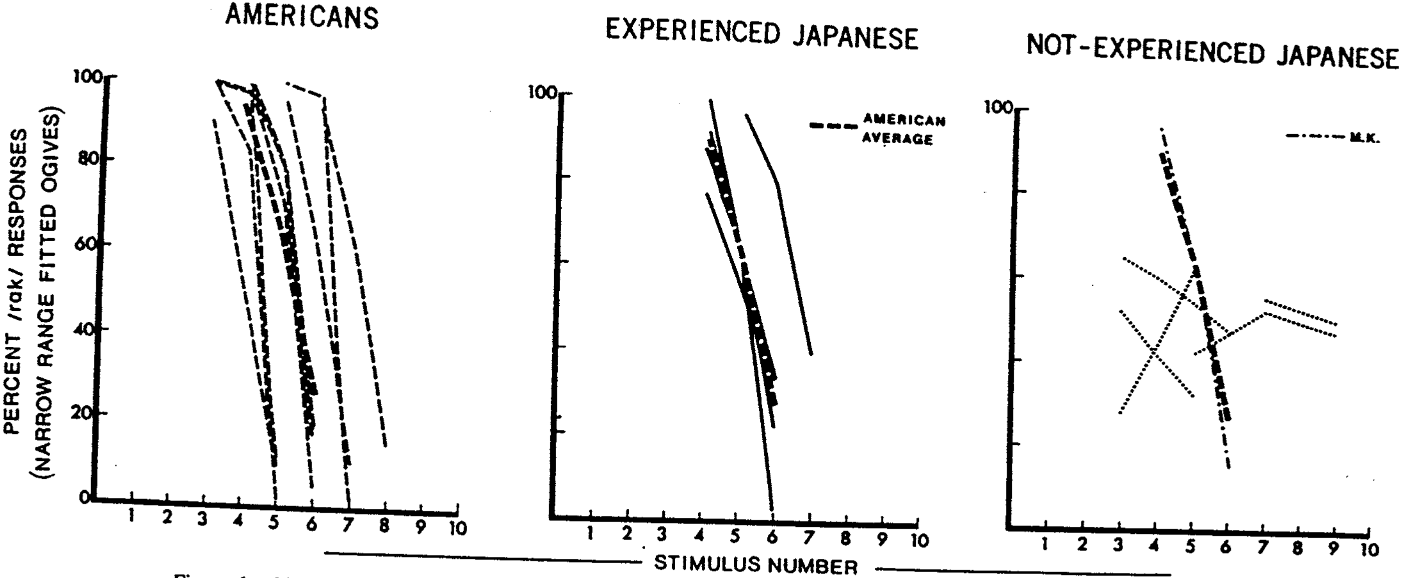

Figure 6. Narrow-range fitted ogives for the category boundaries given by the individual subjects in the three groups.

conversation instruction by native speaker) showed the strongest negative correlation ($r = -.67$, $p < .01$). Factor A (percent of day speaking English in U.S.) showed the smallest correlation ($r = -.33$, ns), and the correlation for Factor C (number of months' experience speaking English with Americans) was –.41, ns. Factor B, which is an indicant of the *intensity* of conversation instruction over an indeterminate period, was more strongly negatively correlated with boundary slopes than was even the total number of hours spent in English conversation instruction (number of hours per week × number of weeks instructed), $r = -.56$, $p < .05$.

An Anomaly: Subject M.K. After completion of the above data analysis, we had the opportunity to test an additional Japanese subject, whose English experience placed him in the Not-experienced group. He was a newly-arrived postdoctoral associate at Yale, and had only been in the U.S. for two weeks at the time of testing. He spoke English less than 25% of the day, and had had no English conversation training by a native speaker, nor was he conversant in any other language besides Japanese. His performance on the three tests, surprisingly, was more similar in many respects to the Experienced Japanese than it was to the Not-experienced group (see Figure 7). His identification function showed a sharp shift near the American boundary (intercept = 5.28), and his discrimination performance was higher even than several of the Experienced Japanese. The major distinction between his data and those of the Experienced Japanese was that both his discrimination functions were bimodal; neither of the peaks fell at his /r/–/l/ boundary, as would have been expected had his ability to discriminate the stimuli been limited in a direct way by his phonetic classifications of them, and as it was for the Americans and the Experienced Japanese.

Relative performances on AXB versus oddity tests. Both discrimination tests were included in our study because of claims that AXB comparisons are less demanding on memory than are oddity comparisons, and provide less of a bias toward phonetic categorization. It has been argued that these circumstances allow subjects to have better access to nonphonetic, precategorical stimulus information under AXB conditions than under oddity conditions (for fuller discussion of this, see Best, Morrongiello, & Robson, 1981 – Experiment 2). These claims lead to the prediction that AXB performance will be better than oddity performance, especially for the Not-experienced Japanese, since they could use nonphonetic auditory memory to aid performance on the AXB test more than on the oddity test. In addition, the oddity boundary-related peak should be sharper than the AXB boundary peak, especially for the Americans and probably for the Experienced Japanese. That is, an auditory memory-related improvement in AXB over oddity performance would affect the within-category judgments more than the between-category judgments.

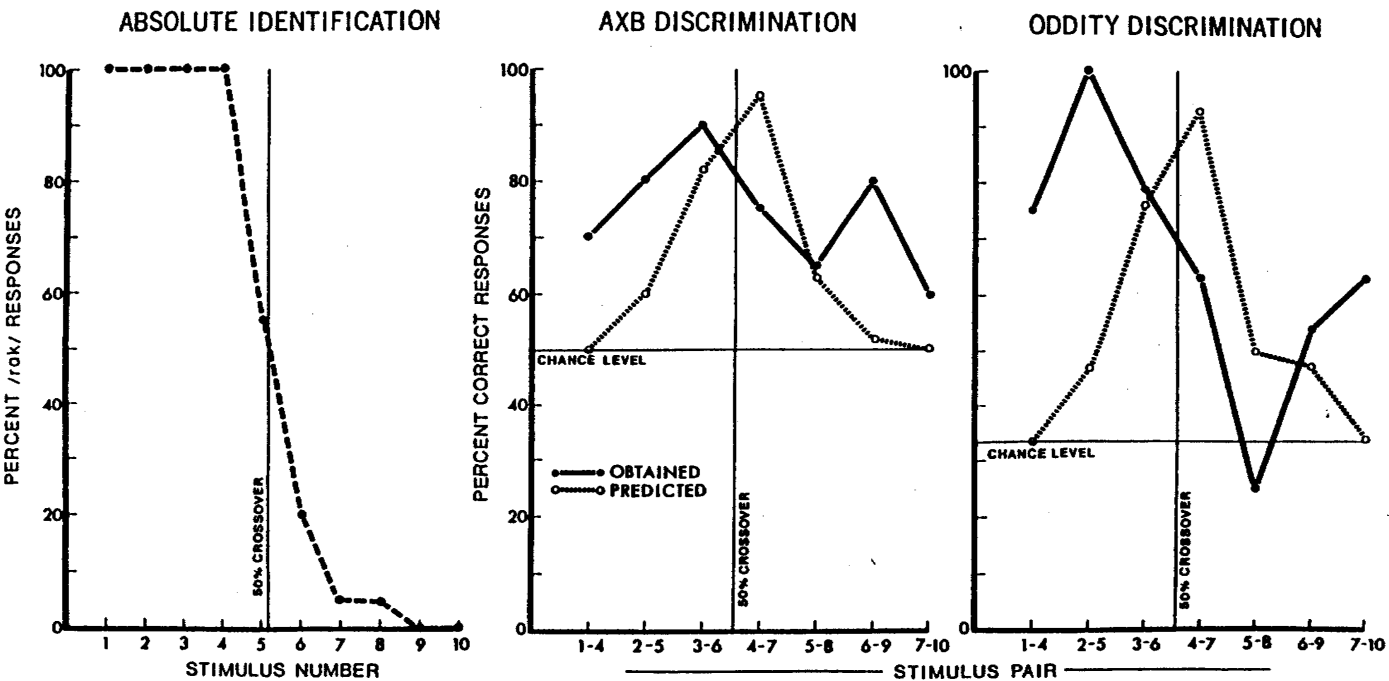

Figure 7. Response functions for subject M. K. on the identification, AXB, and oddity tests.

In order to make a direct AXB-oddity performance comparison, it was necessary to adjust for the difference in chance level performance on the two tests (50% for AXB and 33.3% for oddity). Therefore, performance on the two discrimination tests was recalculated as percentage of above-chance performance. These above-chance performance data were analyzed separately for each group in two-way Test (AXB vs. oddity) × Stimulus Pairs (1–4 through 7–10) ANOVAs.

As can be seen in Figure 8, AXB performance was better than oddity performance for the Americans, according to their significant Test effect, $F(1,4) = 14.90$, $p < .05$. However, the Test × Stimulus Pairs effect for this group did not reach significance, suggesting that, contrary to the auditory memory/phonetic bias predictions, their oddity peak was not consistently sharper than their AXB peak. Their between-category performance was no less affected by the test format than was their within-category performance. Moreover, again in contradiction to the auditory memory/phonetic bias predictions, the Test effect and the Test × Stimulus Pairs interactions failed to reach significance either for the Not-experienced Japanese, or for the Experienced Japanese. It is especially surprising that the oddity discrimination performance of this latter group is closer in form to the ideal picture of categorical discrimination than is their AXB function, in light of suggestions that the oddity paradigm biases subjects toward phonetic categorization rather than discrimination of auditory properties. That is, for these adults, who are learning a nonnative contrast, a bias *toward* phonetic categorization (oddity test) leads to *better* discrimination of between-category comparisons than does a task with a presumably reduced bias toward phonetic categorization and a lower memory demand (AXB).

The sum of the results from the comparison of discrimination tasks does not lend support to the notion that the reason that AXB performance exceeds oddity performance is because the former task allows subjects better access to precategorical (nonphonetic, or auditory) information, at least not when judgments are being made on stimuli whose characteristics approach the acoustic properties found in natural speech. If the auditory memory/phonetic bias picture were correct, all three groups should have fared better on AXB than on oddity judgments. Also, a significant Test × Stimulus Pairs interaction should have been found for the Americans and the Experienced Japanese, indicating that within-category judgments were improved on the AXB task relative to between-category judgments. Furthermore, the Not-experienced Japanese should have shown even greater task effects than the other two groups, since they could use nonphonetic auditory memory but could not rely on phonetic categorization. Instead of the predictions being supported, the pattern of AXB-oddity comparisons across the three groups suggests that performance on *both* tests reflects the effects of phonetic perception. The only group that showed significantly higher AXB than oddity performance was the group that was most experienced with /r/ and /l/ as a phonemic contrast – the Americans. Recall also that this was the

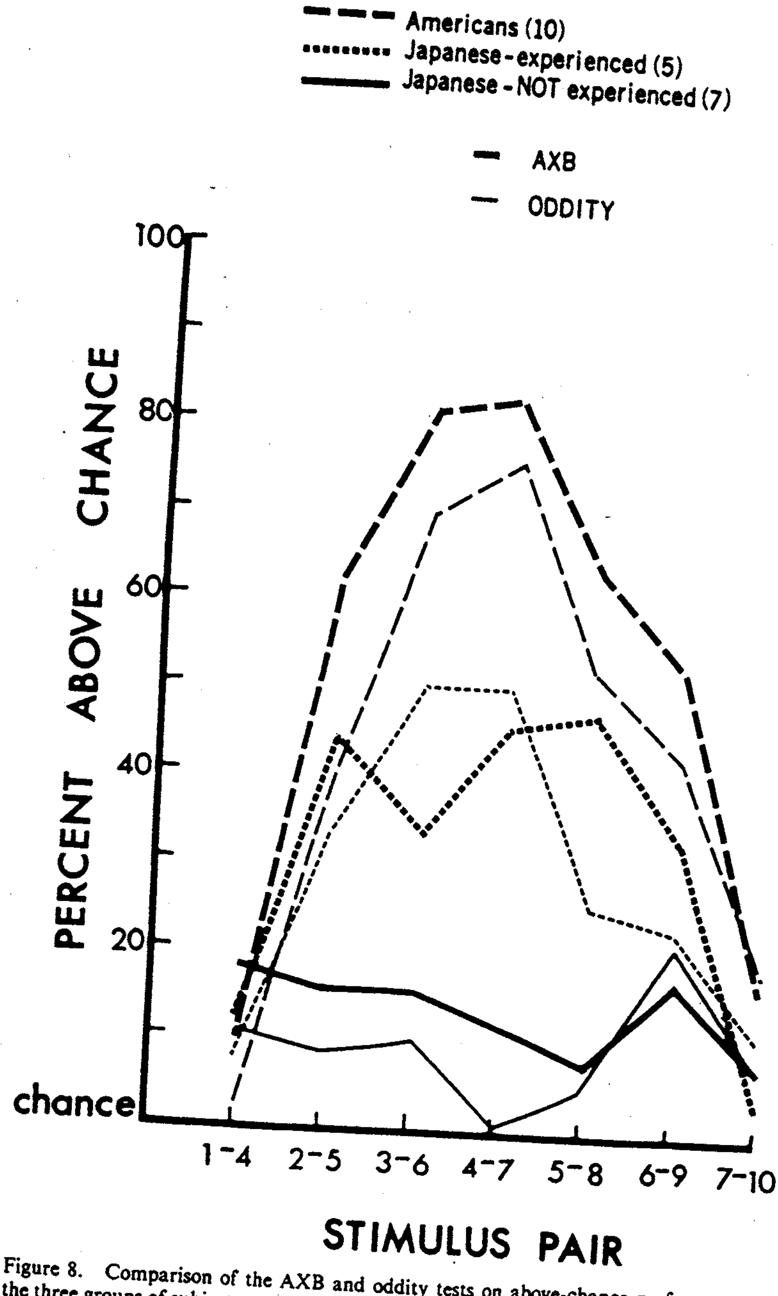

Figure 8. Comparison of the AXB and oddity tests on above-chance performance by the three groups of subjects.

only group whose obtained performance on both discrimination tests was better than predicted by their identification data, and that they only showed better than predicted performance on within-category comparisons that differed in "category goodness." In addition, the nonsignificant Test × Stimulus Pairs interaction for the Americans indicates that their AXB advantage was not due to better accessing of nonphonetic auditory information, but rather that it derived from some improvement in access to specifically phonetic information.

DISCUSSION

This study investigated categorical perception of /r/ and /l/ by native Japanese speakers residing in the U.S. who had had varying amounts of experience in English conversation with native speakers. Bilingual Japanese speakers who were not experienced in English conversation with natives showed, with one exception, near-chance performance across the /r/–/l/ series in the identification task and correspondingly low performance on the discrimination tests. These results corroborate and, for the oddity discrimination test, replicate the earlier Miyawaki et al. (1975) findings with a new stimulus series that provided redundant cues for the phonetic contrast.

The group of focal interest, those bilingual Japanese speakers with relatively intensive conversational experience, performed more similarly to the American-English controls than to their less experienced Japanese counterparts. The ogives derived from the identification function for each of the Experienced Japanese showed a sharp category boundary that was nearly indistinguishable from those of the American-English controls (see Figure 6). Discrimination results were well predicted from the identification data, showing significant peaks in performance at the category boundary for both tests. These results are most encouraging, for they demonstrate that native Japanese speakers learning to converse in English as adults can achieve phonetic categorization of /r/ and /l/ that approximates the categorization behavior of native English speakers.

It is appropriate at this point to discuss the unusually excellent performance of one nonexperienced Japanese subject, M. K. Much to our surprise, his performance on the three tests was more similar to the Experienced Japanese group than to the Not-experienced group (see Figure 7). The major distinction between his data and those of the Experienced Japanese was that the form of his discrimination functions were not predicted by his /r/–/l/ identification results, suggesting that his ability to discriminate the stimuli may not have been directly tied to his phonetic classification of them. However, an alternative explanation to his uncorrelated discrimination responses has not been ruled out. During the identification test, only one stimulus is presented and a categorization response is noted immediately. In contrast, the discrimination tasks require that two or three sounds be held in memory over several seconds before discrimination judgments are made.

Under these memory demands, unstable phonetic representations for these sounds might be disrupted easily, resulting in less consistent performance. We suspect that M.K.'s consistent identification of /r/ and /l/ shows an unusual sensitivity to phonetic distinctions; however, without additional measures of his perceptual behavior, his performance remains an interesting anomaly.

This study has demonstrated that some native Japanese speakers learning English as adults are capable of categorically perceiving /r/ and /l/ in a manner similar to native English speakers. Differences in performance between the Experienced and Not-experienced groups were correlated with differences in conversational experience; however, we cannot rule out a host of variables (e.g., motivation to learn) that might account for differences in performance between the two Japanese groups. Because this study was not designed to test the longitudinal effects of experience, with pre- and post-testing of the same subject on perception of /r/ and /l/, we must infer that the Experienced group represented typical native Japanese speakers, and that they at one time failed to perceive /r/ and /l/ categorically. We are fairly confident that this is the case since each subject was asked about previous problems with /r/ and /l/, and they all reported having great difficulty with this contrast initially, as well as reporting a gradual improvement over time.

The design used here cannot directly answer questions about whether and what kinds of experience produce the change toward categorical perception of phonetic contrasts. Laboratory training studies have had some success in improving /r/–/l/ perception by native speakers of Oriental languages in which the contrast is not phonological. For example, Gillette (1980) reports significant improvement in natural /r/ and /l/ identifications by Japanese and Korean native speakers following several weeks of intensive training with natural speech. Dittmann and Strange (1981) have used a same-different discrimination task with feedback, and produced a change in perception of a synthetic /r/–/l/ series from uniformly poor discrimination to categorical perception by native Japanese speakers.

Future research should be directed toward discovering the perceptual strategies speakers use in their acquisition of this contrast, and determining the conditions that best facilitate acquisition of this contrast by second language learners. Some laboratory training studies currently employ repetitions of minimal pairs of words, natural or synthetic, in listening tasks that require subjects to perform highly differentiated analyses at the level of distinctive features. In accordance with results from first language learners (cf. Menyuk & Menn, 1979), it may be more efficacious, in initial learning of a nonnative contrast by adults, to approximate the first language learning situation in which words are presented in natural speech in sentence contexts and related to objects and events, thus maximizing information at a number of linguistic levels. Following experience with /r/ and /l/ under these conditions, redundant information could be reduced systematically until subjects are required to perform under the most demanding situation, that of making a perceptual distinction between minimal pairs.

ACKNOWLEDGMENTS

Supported by NICHD (NIH) postdoctoral fellowship grant HDO5407 to the first author, NINCDS (NIH) postdoctoral fellowship grant NSO5877 to the second author, NIH grant HDO1994 to Haskins Laboratories, and NIMH grant MH21153 to James J. Jenkins, which supported the third author while she was a visiting scholar at Haskins Laboratories. We'd like to thank Leon Seraphim for help in acquiring Japanese subjects and discussions on Japanese phonetics.

REFERENCES

Abramson, A. S., & Lisker, L. Discriminability along the voicing continuum: Cross language tests. *Proceedings of the 6th International Congress of Phonetic Sciences*. Prague: Academia, 1970.

Abramson, A., & Lisker, L. Voice timing perception in Spanish word-initial stops. *Journal of Phonetics*, 1973, *1*, 1–8.

Best, C. T., Morrongiello, B., & Robson, R. Perceptual equivalence of acoustic cues in speech and nonspeech perception. *Perception & Psychophysics*, 1981, *29*, 191–211.

Dalston, R. M. Acoustic characteristics of English /w,r,l/ spoken correctly by young children and adults. *Journal of the Acoustical Society of America*, 1975, *57*, 462–469.

Dittmann, S., & Strange, W. Effects of discrimination training on the perception of /r/ and /l/ by a native Japanese speaker. Paper presented at the 7th annual *Minnesota Regional Conference on Language and Linguistics*, May 8–9, 1981.

Gillette, S. Contextual variation in the perception of *L* and *R* by Japanese and Korean speakers. *Minnesota Papers in Linguistics and the Philosophy of Language*, 1980, *6*, 59–72.

Glass, G. V., & Stanley, J. C. *Statistical methods in education and psychology*. Englewood Cliffs, N. J.: Prentice-Hall, 1970.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. Perception of the speech code. *Psychological Review*, 1967, *74*, 431–461.

Lisker, L., & Abramson, A. S. A cross-linguistic study of voicing in initial stops: Acoustical measurements. *Word*, 1964, *20*, 384–422.

Lisker, L., & Abramson, A. S. Some effects of context on voice onset time in English stops. *Language and Speech*, 1967, *10*, 1–28.

Lisker, L., & Abramson, A. S. The voicing dimension: Some experiments in comparative phonetics. *Proceedings of the 6th International Congress of Phonetic Sciences*. Prague: Academia, 1970.

Menyuk, P., & Menn, L. Early strategies for the perception and production of words and sounds. In P. Fletcher & M. Garman (Eds.), *Language acquisition*. Cambridge University Press, 1979, 49–70.

Miyawaki, K. *A study of lingual articulation by use of dynamic palatography*. Unpublished masters thesis, University of Tokyo, 1973.

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

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

O'Connor, J. D., Gerstman, L. J., Liberman, A. M., Delattre, P. C., & Cooper, F. S. Acoustic cues for the perception of initial /w,r,l/ in English. *Word*, 1957, *13*, 25–43.

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

Stevens, K. N., Liberman, A. M., Studdert-Kennedy, M., & Öhman, S. E. G. Cross language study of vowel perception. *Language and Speech*, 1969, *12*, 1–23.
Terbeek, D. A cross-language multi-dimensional scaling study of vowel perception. *Working Papers in Phonetics* (University of California at Los Angeles), 1977, *37*.
Williams, L. The voicing contrast in Spanish. *Journal of Phonetics*, 1977, *5*, 169–184.

APPENDIX A

Nominal parameter values for the /rak/-/lak/ stimulus series. 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 ms into the syllable (T – 35).

Stimulus number	*F1 duration (ms)*		*Formant center 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 portion of stimuli: formant center frequencies (in Hz)

	Vowel			*Final closure*		
F1 start	*F1*	*F2*	*F3*	*F1*	*F2*	*F3*
349	621–707	1198–1233	2557	621	1288	2104

Revision received: 18 September 1981