Wode, H. 1978. The beginnings of non-school room L2 phonological acquisition. *International Review of Applied Linguistics* 16:109–25.
Wode, H. 1981. *Learning a Second Language. I. An Integrated View of Language Acquisition.* Tübingen: G. Narr.
Yamada, R. A., and Tohkura, Y. 1990. Perception and production of syllable-initial English /r/ and /l/ by native speakers of Japanese. *Proceedings of the International Conference on Spoken Language Processing.*
Yeni-Komshian, G. 1993 (personal communication)

Chapter • 10

Age and Acquisition of Second Language Speech Sounds

Perception of American English /ɹ/ and /l/ by Native Speakers of Japanese

Reiko A. Yamada

Mina is a 6-year-old girl who has just returned to Japan from a 1-year stay in the United States. Her speech is being recorded for an experiment. "Count from 0 to 10," says the experimenter. "Zero, one, TSu, Sree," Mina responds. She pronounces the English words with a Japanese accent. It seems that a 1-year stay in the United States may be insufficient time for Japanese children to acquire English sounds. "Say the days of the week," says the experimenter. "Sunday, Monday, Tuesday," Mina responds. Surprisingly, Mina now produces fluent and unaccented American English speech. Mina's results on a perception test of the English /ɹ/–/l/ contrast, which is difficult for Japanese to acquire, shows quite consistent identification, similar to that of American children of the same age as Mina (see the Appendix). It seems that both the experience of staying in the United States for 1 year and exposure to Japanese-accented English in Japan affected Mina's perception and production of English sounds. When she is retested 18 months later, it is shown that she retained her /ɹ/–/l/ production and perception abilities, although she has not been exposed to an English-speaking environment during these 18 months.

Language usage is one of the remarkable species-specific characteristics of *Homo sapiens*. There are thousands of languages, each with its own phonemic system. Although almost everyone acquires his or her native language with seeming ease, learners often have difficulties in acquiring a new language and its phonemic system, especially in adulthood. On the other hand, multilingualism does exist, which demonstrates that humans are able to acquire and use multiple language systems.

For both humans and other animals, learning is shaped by biological predispositions and capabilities (cf. Seligman and Hager 1972). Language learning is no exception, and it is purported to have various biological foundations and limitations (cf. Lenneberg 1967). These biological factors have been studied through behavioral and physiological experiments, under more controlled conditions in animals than in humans, because, in animal studies, methodological and ethical restrictions are less severe than in human studies. In such animal-learning studies, song learning by songbirds has several characteristics also seen in human language acquisition; for example, similarity between babbling in humans and subsong in songbirds; cerebral lateralization of vocal control; and existence of local dialects (cf. Kroodsma and Miller 1982). Regarding the biological limitations for song-learning, there are limitations in the age of learning, and in learned objects. Many species of songbirds can learn their songs only during limited periods; that is, *critical periods*. Zebra finches subsequently acquire their song if they are exposed to the adult model between the 40th and the 80th day after hatching (Immelmann 1969). White-crowned sparrows imitate models presented during the first 50 days after hatching (Marler 1970). Furthermore, the period of template shaping (perceptual learning) and the period of practice of vocalizations (motor learning) are often separated. In white-crowned sparrows, practice using the acquired template with auditory feedback begins 3 or 4 months after the termination of the sensitive period for perceptual learning (Marler 1970). In addition to this time-domain limitation, selectivity of imitated models is also observed in many songbirds. If reared in isolated circumstances during infancy, white-crowned sparrows cannot learn species-typical songs and sing a simple pattern. However, if both their own species' song and another species' song are present in their isolated environment, they will learn to sing only their species' song (Marler 1984).

In humans, the results from many cross-language speech perception studies suggest that native language phoneme perception is an acquired ability. Some of the strongest support for this conclusion comes from the performance of native speakers of Japanese on American English (AE) /ɹ/–/l/ perception and production tests. However, if Japanese are exposed to an English-speaking environment at a young age, as Mina was in the opening of this chapter, they acquire nearly native-like speech perception and production abilities in English. Thus, there appears to be a strong age effect on second-language speech acquisition. (See also Flege this volume.)

The critical period hypothesis in native language acquisition has been supported by a study in which the age of recovery from acquired aphasia was examined; subjects older than 8 years have a much lowered probability of reacquiring their native language than younger subjects (Lenneberg 1967). Recent cross-language studies also provide evidence that there is an age effect in acquiring non-native phonemes. Older learners tend to fail in acquiring second language speech perception/production more often than young learners (Flege 1990, this volume; Mack 1989; Oyama 1976; Tahta et al. 1981a, 1981b). Regarding English /ɹ/ and /l/ perception and production by native speakers of Japanese, it has been shown that those who had the experience of residing in the United States in childhood showed native-like characteristics; whereas, those who had similar experiences at an older age did not (Cochrane 1980; Yamada and Tohkura 1992). Needless to say, critical/sensitive periods themselves cannot explain the phenomena, but the mechanisms underlying them must be studied. From this standpoint, the nature of the relationship between age and the acquisition of second language is worth studying further.

In this paper I address the question of what effect age has on subjects' ability to acquire non-native phonemes; specificially how the starting age and duration of exposure to an English-speaking environment affects the acquisition of /ɹ/ and /l/ perception by native speakers of Japanese.

EXPERIMENT

Method

Three kinds of perception tests were conducted: 1) identification and discrimination tests of a synthesized /ɹ–l/ continuum using stimuli in which multiple acoustic cues to the contrast were varied concurrently; 2) identification of synthesized stimuli whose F2 and F3 (second and third formants) onset and transition frequencies and F1 (first formant) transition duration varied factorially; and 3) identification of natural words containing AE /ɹ/, /l/, and /w/.

The first test was conducted in order to examine the categorical perception of a synthetic /ɹ–l/ series. A pair of English words, "right" and "light," was selected as the minimal pair contrasting initial consonants, and a synthetic /ɹait/–/lait/ continuum was generated with Klatt cascade formant synthesis (Klatt 1980). The continuum consisted of 17 stimuli (ST1–ST17) in which three acoustic parameters, F2 and F3 onset and transition frequencies and F1 transition duration, were varied (Fig.1; see Yamada and Tohkura 1992 for details). The stimuli were synthesized and reproduced through 16-bit digital–analog conversion at a sampling frequency of 20kHz and low-pass filtering with a cutoff frequency of 10kHz. The experiment included an identification test

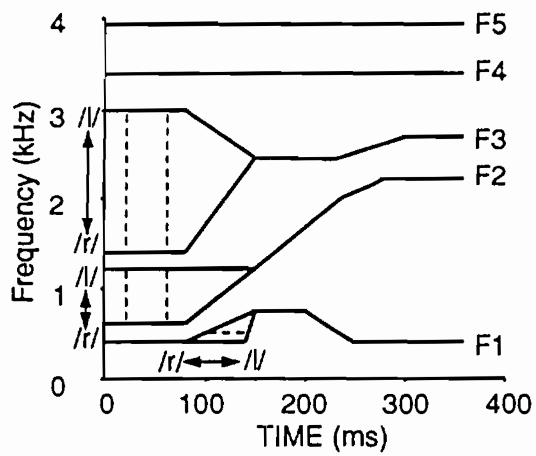

Figure 1. Schematic spectrographic representation of synthetic "right"–"light" stimulus continuum (final /t/ release not shown).

(Test 1a) and a 4-step ABX discrimination test (Test 1b). In the identification test, subjects were asked to identify the initial consonant, and to make a forced choice among /ɹ/, /l/, and /w/ by checking a corresponding response category on an answer sheet. Each subject identified each stimulus six times. In the ABX discrimination test, 13 paired comparisons were presented in triplets. In each triplet, the two stimuli to be compared were presented first (A and B), and then either A or B was presented again (X). Subjects judged whether the third stimulus (X) was identical to the first one (A) or the second one (B) in a forced-choice task. All the stimulus pairs with four steps between; that is, ST1 versus ST5, ST2 versus ST6 . . ., ST13 versus ST17, were compared. Subjects made eight judgments of each comparison pair, among which the order of A, B, and X were counterbalanced; each ABA, ABB, BAA, and BAB, appeared twice. In a trial, the members of each triplet were separated by a 1 second ISI; there were 3 seconds between triplets. Several sessions with different stimulus randomizations were recorded on a digital audiotape using a DAT recorder (SONY DTC-1000ES). In the test, stimuli were presented to the subjects through headphones (STAX SR Lambda pro) binaurally.

The second test (Test 2) was conducted to study the usage of the spectral and temporal cues for the contrast in more detail. A synthesized /ɹait/–/lait/ series similar to the first test was used, except that the F2 spectral parameter varied independently from the F3 spectral and F1 temporal parameters. Onset frequency of F2 varied from 800Hz to 1400Hz in 200Hz steps; F3 varied from 1200Hz to 3000Hz in 200Hz steps. First formant transition duration varied from 70ms to 16ms in 6ms steps as the F3 frequency varied from 1200Hz to 3000Hz.

There were 37 stimuli in total resulting from the pairing of four F2 conditions and ten F3/F1 conditions excluding three contradictory pairs; for example, pairs where F2 onset frequency was greater than or equal to F3 onset frequency. Subjects were asked to identify the initial consonant in a way identical to that in the identification Test 1a. All the subjects judged each stimulus six times.

The third test (Test 3) was an identification test of word-initial /ɹ/, /l/, and /w/ in natural words. The stimuli and procedure were those used in Yamada and Tohkura (1992). Sixteen triplets of words (a few of them were nonwords), in which only the first consonant differed, were used as materials. Two native speakers of American English (AE), one male and one female, uttered the 16 triplets, making 96 stimuli in total. Subjects were asked to identify the initial consonant of each stimulus and forced to choose /ɹ/, /l/, or /w/. All the subjects judged all 96 stimuli.

Subjects

Forty-two native speakers of AE (AE subjects), 120 native speakers of Japanese who had never lived abroad (J subjects), and 156 native speakers of Japanese who had resided in the United States (JE subjects) served as subjects. Criteria for participation in the experiment for each subject group are shown in Table 1a. The duration of the stay in the United States of JE subjects varied from 1 to 23 years (Table 1b). This group was divided into subgroups according to the number of years in the United States (see Table 1b).

Results

Categorical Perception (Test 1a and 1b). Figure 2 presents the results of the first test pooled across subjects in each language group, A, J, and JE. Identification rates for each response category are shown in the top panels, and accuracy in the ABX discrimination test is shown in the lower panels. (Chance level for the discrimination test is 50%.) Subjects in Group A consistently (≥ 90%) identified five stimuli (ST1–5) as /ɹ/ and seven stimuli (ST11–17) as /l/. A few /w/ responses appeared for ST1–11 and peaked at ST7–ST8, but with a maximum identification rate of only 38%. The ABX discrimination function for Group A peaked around the ST7-ST11 pair. In contrast, Group J showed inconsistent identification, but tended to identify four or five /ɹ/-side stimuli (ST1–5) as /ɹ/ most frequently, seven /l/-side stimuli (ST11–17) as /l/ most frequently, and intermediate stimuli (ST6–9) as /w/ most frequently. In addition, they showed many /w/ responses for /ɹ/-side stimuli (ST1–5), to which Group A listeners rarely responded /w/. The function for ABX discrimination was flat, and only

Table I. Subjects used in the experiment

a: Criteria for subject groups

AE

1. Native speaker of A
2. Born and raised in U.S.A
3. Parents are native speakers of AE
4. Living in Japan less than 2 years

J

1. Native speaker of Japanese
2. Parents are native speakers of Japanese
3. Never lived abroad

JE

1. Native speaker of Japanese
2. Once lived in the continental U.S.A. for more than 1 year
3. Never lived in a foreign country other than the U.S.A.
4. Spoke AE at all times at school or in business while in U.S.A.
5. Went to school or business at least 5 days a week while in U.S.A.
6. Received no special training for speaking AE in Japan.

b: Duration of the stay in U.S.A., no. of subjects, mean duration of the stay, mean starting age, and mean age at the time of this experiment for each JE subgroup, JE1 to JE8, and of JE

Group	Duration (range)	No.	Mean yrs.	Starting age	Age at testing
JE1	1 year	46	1	15.3	20.9
JE23	2–3 years	33	2.5	11.7	19.5
JE45	4–5 years	34	4.6	9.6	19.4
JE67	6–7 years	31	6.4	6.7	18.2
JE8	8–23 years	12	12.0	5.4	19.6
JE(all)	1–23 years	156	4.0	10.2	19.6

slightly higher than the 50% chance level. These results from AE listeners and those from Japanese listeners replicated previous reports (Yamada and Tohkura 1991, 1992).

Results pooled across all the subjects in the Group JE showed an intermediate pattern between Groups A and J. On average, JE subjects identified ST1–ST7 as /ɹ/ and ST11–ST17 as /l/ at lower rates than Group A but at higher rates than Group J. The /w/ responses for /ɹ/-side stimuli also showed an intermediate rate between Group A and J. The ABX discrimination for the JE group was also intermediate to A and J groups, showing a peak of most accurate discrimination around pair ST 7–ST11, which was somewhat lower than the A group.

Use of Cues (Test 2). Figure 3 shows the results from Test 2. This test replicates the procedures used previously in Yamada and Tohkura (submitted) but with a reduced set of stimuli. In the previous

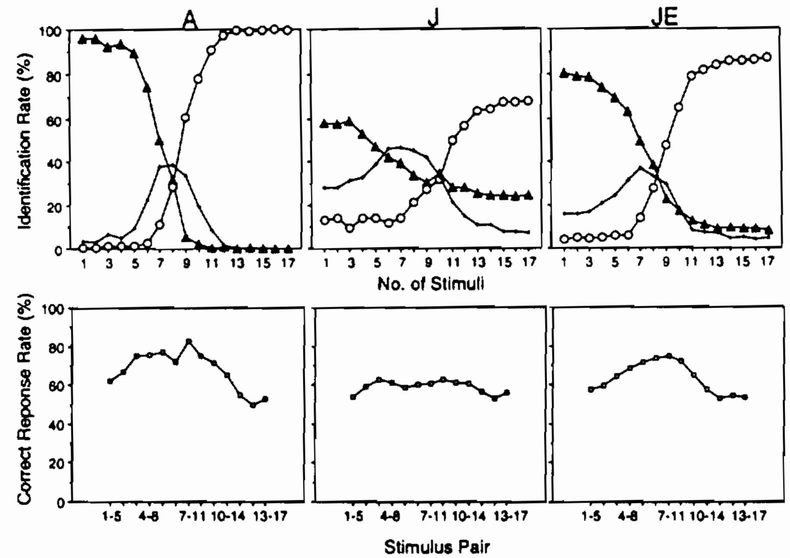

Figure 2. Identification (above) and ABX discrimination (below) of the synthetic /ɹ–l/ continuum by Group A (left), Group J (middle) and Group JE (right).

study, it was shown that when listeners were asked to identify stimuli whose F2 frequency and F3 frequency/F1 transitional duration varied factorially, listeners used both F2 and F3/F1 cues, while AE listeners used F3/F1 cues as the predominant cues to distinguish /ɹ/ and /l/. Figure 3 shows the identification pattern on an F2–F3/F1 plane for each language group: Bold lines separate the plane into three or four regions according to the most frequent responses for each stimulus; for example, subjects responded to the stimuli in the /ɹ/ region most frequently as /ɹ/ among /ɹ/,/l/ and /w/ choices. The response rates of these modal choices for each stimulus are also represented by the grey pattern, with darker grey patterns showing a higher response rate. The results replicated the difference in usage of cues between A and J subjects found in the previous experiment (Yamada and Tohkura 1990, submitted): Group A subjects generally identified the stimuli whose F3 onset frequency was higher than 2.0kHz as /l/ most frequently, the other stimuli as /ɹ/. They identified only one stimulus whose F2 frequency was low and F3 frequency was around 2.0kHz as /w/. In contrast, Group J's identification pattern was more affected by F2 frequency than was that of Group A. They tended to identify the stimuli as /l/ not only depending on the F3/F1 value but also on F2 frequency value. They identified the stimuli with high F2 frequencies

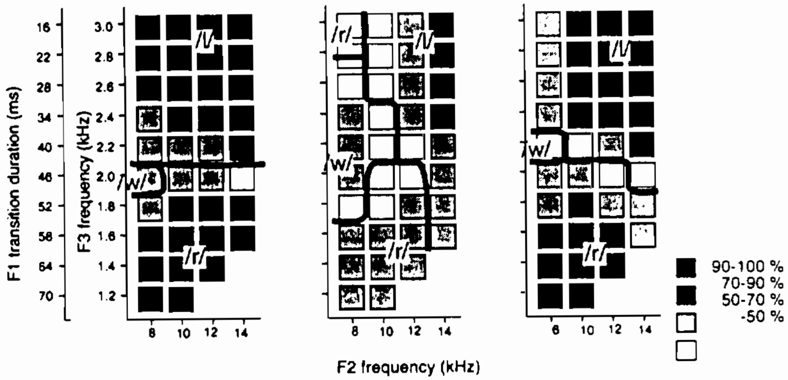

Figure 3. Identification of synthetic stimuli on F2/F3 plane by Group A (left), Group J (middle) and Group JE (right).

as /l/. In addition, Group J subjects had a much larger /w/ response area than Group A. Group JE subjects again showed an intermediate pattern between A and J. They seemed to distinguish /ɹ/ and /l/ mainly, but not solely, using the F3/F1 cue; their pattern was affected by F2 frequency when F3/F1 cues were ambiguous. Response consistency was lower than Group A, but higher than Group J.

Identification of Natural Tokens (Test 3). The correct response rate for natural tokens in Test 3 were 98.0% for Group A, 66.6% for Group J, and 84.6% for Group JE, as shown in the right-hand column of Table 2. Again, Group JE showed an intermediate score between Group A and J.

Effect of Age. From the above results, it is clear that the three language groups, A, J, and JE, behaved differently on all tests. The difference between perception performance by Groups A and J replicate those frequently reported in previous research (Miyawaki et al. 1975; Mackain et al. 1981; Mochizuki 1981, Yamada and Tohkura 1992). Here, the pattern for Group JE is of most interest. The fact that Group JE behaved differently from both Group A and J suggests that the experience of living in the United States affected perception by the Japanese speakers. In other words, these results demonstrate that Japanese speakers' perception of /ɹ/ and /l/ changes significantly through their exposure to an AE-speaking environment. However, this change did not occur to the same extent for all subjects. Some adapted their perception to the AE-speaking environment, while others did not. In fact, there were large individual differences in the present results; the performance of some of the JE subjects showed an identical pattern to Group A subjects, while some showed the typical pattern of Group J,

and others showed an intermediate pattern. As a result, the averaged pattern of Group JE is intermediate between Group A and Group J.

One of the strongest reasons for these individual differences within JE was hypothesized to be the age of the subjects when they were exposed to the AE-speaking environment. To investigate this effect of age on the results of perception tests, correlations between experience variables and perception variables were computed. Canonical correlation analysis was performed (Hotelling 1935, 1936) in which the relationship between two sets of variables (each set being able to contain multiple variables) was investigated. Two age-related variables, OA (onset age) and DUR (duration of the stay), were used as experience variables. The results of the four perception tests were used as perception variables. In Test 1A, /ɹ/-response rates for ST1–ST8 and /l/-response rates for the rest of the continuum ST9–ST17, were averaged; in Test1B, the averaged discrimination rates for all pairs was used. In Test 2, /ɹ/-response rates for the stimuli whose F3 value was lower than 2.0kHz and /l/-response rates for the stimuli whose F3 value was higher than 2.2kHz were averaged. In Test 3, averaged correct response rate was used (Table 2). The first canonical correlation was 0.589, which was larger than any of the correlations among the original variables (DUR and TEST 2 showed largest correlation of 0.486). The standardized first canonical coefficients for the experience variables were -0.597 for OA and 0.572 for DUR. The correlations between OA and DUR and the first canonical variable were -0.862 and 0.849. The correlations between experience factors and the canonical variables of perception were -0.507 for OA and 0.499 for DUR. These results are interpreted as showing that both OA and DUR contributed to the experience variable; later OA reduces, and longer DUR increases accuracy in the perception tests to approximately the same extent. In the present data, there was a strong tendency that the subjects with earlier OA also had longer DUR. (The correlation between OA and DUR was -0.465; this value is about the same as the correlation between canonical variables of perception and OA [-0.507] and DUR [0.499].) As a result, it can be concluded that age-related factors affect the /ɹ/–/l/ perception in the way that longer DUR and/or earlier OA facilitates the Japanese listeners' perception ability of English /ɹ/ and /l/. However, the interaction between the effect of OA and DUR is not clear.

To investigate the effect of OA in the present data while avoiding the correlation between OA and DUR, JE subjects were classified into subgroups depending on their DUR values; JE1 (DUR = 1 year), JE23 (DUR = 2 or 3 years), JE45 (DUR = 4 or 5 years), JE67 (DUR = 6 or 7 years), and JE8 (DUR ≥ 8 years). The mean scores on the four perception tests for each of these subgroups, which were the same values used in the above canonical analysis, are shown in Table 3. In addition,

Table 2. Score from Test 1a, 1b, 2, and 3 of Subject Group A, J, JE, and JE Subgroups.

Group	Test 1a	Test 1b	Test 2	Test 3
A	85.2	67.9	84.5	98.0
J	52.1	58.8	47.6	66.6
JE	72.9	63.1	71.3	84.6

See text for derivation of scores.

Table 3. Score from Test 1a, 1b, 2, and 3 of Five Subgroups of Group Group JE; JE1, JE23, JE45, JE67, and JE8.

Group	Test 1a	Test 1b	Test 2	Test 3
JE1	66.9	61.3	61.2	77.0
JE23	65.6	63.4	65.6	79.0
JE45	74.4	61.3	74.3	88.8
JE67	81.9	66.1	82.5	92.9
JE8	88.0	66.7	85.9	95.2

See text for derivation of scores.

these data are plotted in Figure 4, with the data from Groups A and J. In Figure 4, distributions are shown in each panel; the distribution of Group A with high scores, the distribution of Group J with low scores, and the distribution of Group JE subgroups shown with bold lines.

An inspection of the data in Figure 4 suggests that the five subgroups can be classified into three groups with respect to how their performance relates to that of J and A groups: a) JE1, JE23; b) JE45, JE67; c) JE8. JE1 and JE23 are similar to J group subjects. The JE8 subjects performed almost exactly the same as A subjects. The JE45 and JE67 groups were intermediate. Note, in particular, that there seem to be significant differences in the shape of the distributions in the Test 1a–1b plane.

It is observed that the distribution in the JE subgroups with shorter duration of experience tend to overlap with that of Group J, and subgroups with longer duration of experience overlap with the distribution of Group A. As noted above, this is not a result of duration alone, but could be a joint effect of duration and starting age. The effect of OA in each of the duration subgroups was analyzed with the exception of Group JE8. (For Group JE8, 10 of 12 subjects started their living experiences in the United States before 7 years of age.) In each subgroup, the subjects were divided into groups according to their OA: 0–7 years, 8–15 years, and 16–25 years. The results of Test 1a, Test 2, and Test 3 were compared for these OA groups. As shown in Figure 5, the mean value showed a tendency for the group with earliest OA to show higher scores in the perception tests although these differences

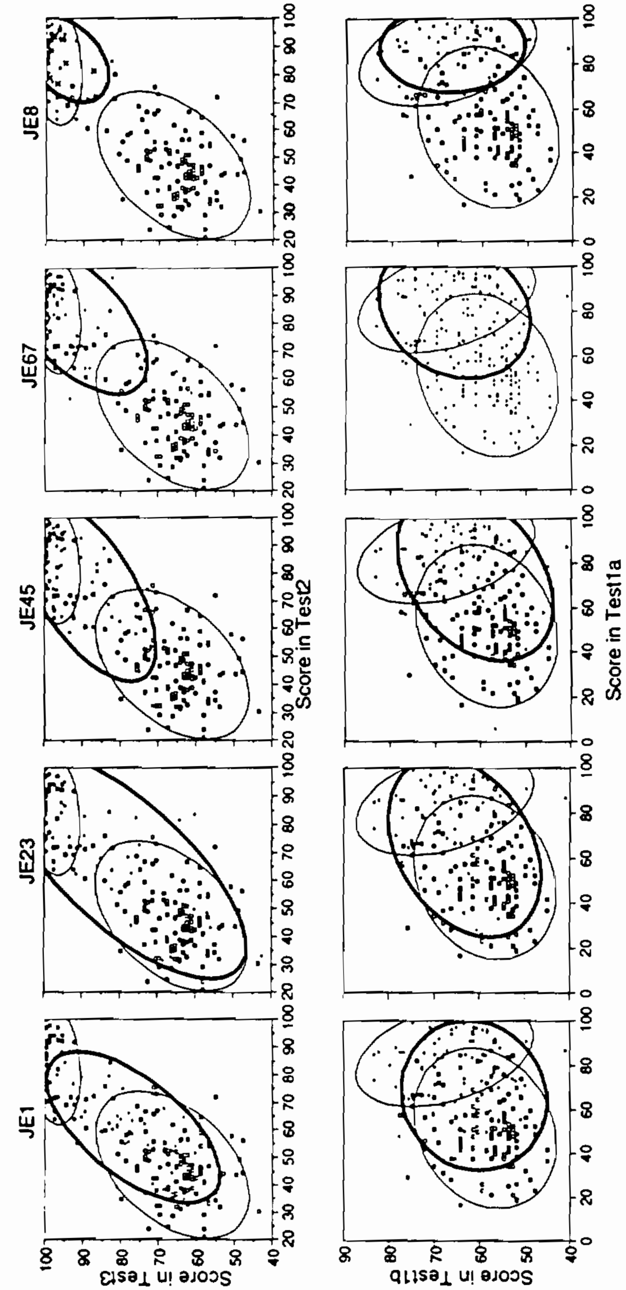

Figure 4. Results of the four perception tests in each JE subgroup (bold-faced ovals). The data for Group J and those for Group A are also plotted in each panel.

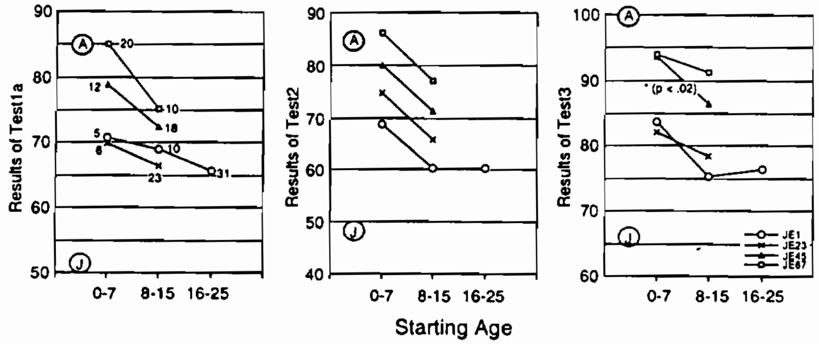

Figure 5. The effect of starting age of the experiences.

were not significant when compared by an unpaired *t*-test, except for the subgroup JE45 in Test 3 (the scores of those whose starting age was 0–7 and that of those of 8–15 differed significantly; $p < .02$).

DISCUSSION

The results described above showed that experience living in the United States significantly affects the perception of /ɹ/ and /l/ by native Japanese speakers. The younger the speakers and the longer the duration of their experience, the better their ability to perceive /ɹ/ and /l/. However, the interaction between starting age and duration of experience could not be analyzed with the present data, because these two variables were confounded. It is supposed that both age and experience are significant factors in determining L2 learning abilities. These can be further classified into two types of biological and learning factors. Biological factors affect the acquisition of non-native sounds by decreasing the plasticity in neural systems with aging. Learning factors include various interactions with circumstances, and at least the following factors could have affected the present result.

First, the initial state of JE subjects when they were first exposed to an English speaking environment was marked by large individual differences, and these differences may have affected the present results. Note that Group J's data showed large individual differences. These differences might have resulted from differences in English education in Japan. English education usually starts in junior high school at the age of about 12 years old and lasts until graduation from high school or college. It is difficult to acquire English sounds through this ordinary English education, because it is biased to grammatical instruction, and more importantly, there is little necessity to use English in daily life in Japan. However, improvement in the ability to distinguish /ɹ/ and /l/ during this education differs much between individuals, maybe

depending on a student's motivation and/or ability to learn English and also on the quality of English lessons received. These large individual differences were also reported in Yamada and Tohkura (1992). The differences in the initial state must be considered when living experiences in a new language environment are manipulated in the study, especially when subjects learn the language before they begin their residency. However, it is almost impossible to control or describe the initial state in a cross-sectional design as in the present study.

Second, the quality of L2 experience may have affected the results, although it is difficult to control or measure. The JE subjects used English in their daily life at least 5 days a week during their stay in United States. However, the extent of their use of English may have differed among subjects. Third, the English education in Japan might lead to qualitative differences in listeners' strategy when distinguishing /ɹ/ and /l/ sounds. When Japanese listeners who have English knowledge but cannot distinguish /ɹ/ and /l/ by ear, try to distinguish these phones, they may rely on context rather than the acoustic signal itself. In a conversational situation, sentences, words, and so forth, contain context which helped listeners to discriminate /ɹ/ and /l/. In fact, it was experimentally revealed that Japanese listeners' identification of words with /ɹ/ and /l/ are significantly affected by the familiarity of the words (Yamada et al. in press).

It is necessary to collect more controlled samples in order to have a clearer view of the effect of age in acquiring new phonetic contrasts through living in a foreign-language-speaking environment. Other approaches are also necessary; for example, a longitudinal study in which the ability before, during, and after the stay in the United States of specific subjects is traced. Another effective way is a training study, in which subjects' age is manipulated. Strange and Dittmann (1984) trained Japanese subjects to distinguish /ɹ/ and /l/ perceptually using a discrimination task and synthesized stimuli. The results did not transfer to novel stimuli in a generalization test. However, recent laboratory training studies of /ɹ/–/l/ using stimuli produced by several talkers in which the contrast occurs in various vocalic contexts were effective in training Japanese speaking adults to identify English /ɹ/ and /l/ (Logan et al. 1991; Lively et al. 1991; Pisoni and Lively this volume).

As with many other learning phenomena, when acquiring non-native phonemes, learners must adapt to the given circumstances by trying to acquire a new behavior or to reallocate their own structures of behaviors. On the other hand, in acquiring a new behavior, acquisition of new mental functions or reorganization of such functions may be indispensable. In the training experiment, we must consider that many things such as training task factors (structure of training procedures and stimuli used in the training), linguistic factors (language which the learner has already acquired), and social and cultural factors (the education system, necessity to use the language, etc.) must be considered. By manipulating these factors effectively in the laboratory training experiment, even adults have a chance to acquire non-native phonemes smoothly. Such training task factors are being studied through laboratory training experiments (Logan et al. 1991; Strange 1991; Lively et al. 1994; Logan and Pruitt this volume; Pisoni and Lively this volume). With respect to linguistic factors, facilitative or interference effects related to the L1 phonetic system were reported and a couple of hypotheses on such effects were proposed (see Best this volume; Best et al. 1988, Best and Strange, 1992; Flege this volume; and Flege 1990). However, in non-laboratory studies, adaptation behavior should also be considered. If other competing or supplementary

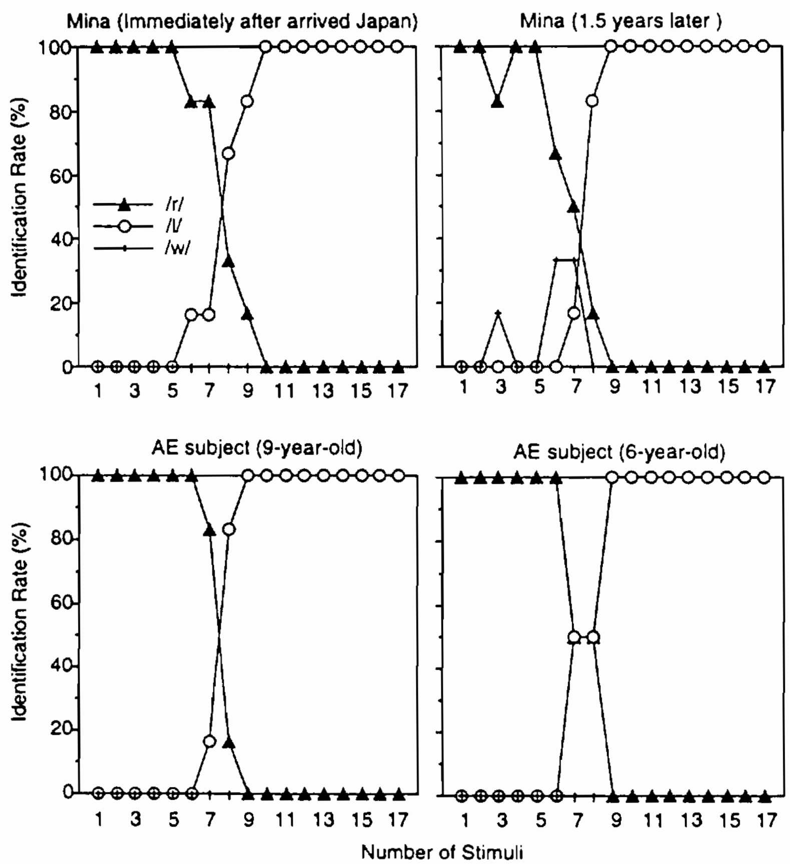

Figure A. Identification data for Mina at two time intervals (above), with data for same-age native English control subjects (below).

adaptive behaviors (such as understanding through contexts, gestures, and visual cues) are acquired, learners may suppress phoneme perception learning itself. Finally, experimental training studies must be arranged that investigate the effect of age of the subject on the success of training procedures and the rates of perceptual improvement. In this way, the correlated variables found in field experiments may be disentangled.

REFERENCES

Best, C. T., McRoberts, G., and Sithole, N. 1988. Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14:345–60.

Best, C. T., and Strange, W. 1992. Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics* 20:305–30.

Cochrane, R. M. 1980. The acquisition of /r/ and /l/ by Japanese children and adults learning English as a second language. *Journal of Multilingual and Multicultural Development* 1:331–60.

Flege, J. E. 1990. Perception and production: The relevance of phonetic input to L2 phonological learning. In *Crosscurrents in Second Language Acquisition and Linguistic Theories*, ed. C. Ferguson and R. Huebner. Philadelphia: John Benjamins.

Hotelling, H. 1935. The most predictable criterion. *Journal of Educational Psychology* 26:139–42.

Hotelling, H. 1936. Relations between two sets of variables. *Biometrika* 28:321–77.

Immelmann, K. 1969. Song development in the zebra finch and other estrildid finches. In *Bird Vocalizations*, ed. R. A. Hinde. Cambridge: Cambridge University Press.

Klatt, D. H. 1980. Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America* 70 1261–71.

Kroodsma, D. E., and Miler, E. H. 1982. Song learning and its consequences. In *Acoustic Communication in Birds* (vol. 2). New York:Academic Press.

Lenneberg, E. H. 1967. *Biological Foundations of Language.* New York: Wiley, 1967.

Lively, S. E., Pisoni, D. B., and Logan, J. S. 1991. Some effects of training Japanese listeners to identify English /r/ and /l/. In *Speech Perception, Production, and Linguistic Structure*, ed. Y. Tohkura, E. Vatikiotis-Bateson, and Y. Sagisaka. Ohmsha / Amsterdam: IOS Press.

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., and Yamada, T. Training Japanese listeners to identify English /r/ and /l/ III Long-term retention of new phonetic categories. *Journal of Acoustical Society of America* 96:2076–87.

Logan, J. S., Lively, S. E., and Pisoni, D. B. 1991. Training Japanese listeners to identify /r/ and /l/. *Journal of the Acoustical Society of America* 89:874–86.

Mack, M. 1989. Consonant and vowel perception and production: Early English–French bilinguals and English monolinguals. *Perception & Psychophysics* 46:187–200.

Mackain, K. S., Best, C. T., and Strange, W.. 1981. Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics* 2:369–90.

Marler, P. 1970. A comparative approach to vocal learning: Song development in white-crowned sparrows. *Journal of Comparative and Physiological Psychology* 71:1–25.
Marler, P. 1984. Song learning: Innate species differences in the learning process. In *The Biology of Learning*, ed. P. Marler and H. S. Terrace. Berlin: Springer-Verlag.
Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., and Fujimura, O. 1975. An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics* 18:331–40.
Mochizuki, M. 1981. The identification of /r/ and /l/ in natural and synthesized speech. *Journal of Phonetics* 9:283–303.
Oyama, S. 1976. A sensitive period for acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research* 5:261–83.
SEligman, M. E. P., and Hager, J. L. 1972. *The Biological Boundaries of Learning*. New York: Appleton-Century-Crofts.
Strange, W. 1991. Learning non-native phoneme contrasts: Interactions among subject, stimulus and task variables. In *Speech Perception, Production and Linguistic Structure*, ed. Y. Tohkura, E. Vatikiotis-Bateson, and Y. Sagisaka. Ohmsha / Amsterdam: IOS Press.
Strange, W., and Dittmann, S. 1984. Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Perception & Psychophysics* 36:131–45.
Tahta, S., Wood, M., and Loewenthal, K. 1981. Age changes in the ability to replicate foreign pronunciation and intonation. *Language and Speech* 24:363–72.
Tahta, S., Wood, M., and Loewenthal, K. 1981. Foreign accents: Factors relating to transfer of accent from the first language to a second language. *Language and Speech* 24:265–72.
Yamada, R. A., and Tohkura, Y. submitted. Perceptual cues of English /r,l,w/ for native speakers of American English and those for native speakers of Japanese. *Perception & Psychophysics*.
Yamada, R. A., and Tohkura, Y. 1992. The effects of experimental variables on the perception of American English /r/ and /l/ by Japanese listeners. *Perception & Psychophysics* 52:376–92.
Yamada, R. A., Kobayashi, N., and Tohkura, Y. in press. Effect of word familiarity on non-native phoneme perception: Identification of English /r/, /l/, and /w/ by native speakers of Japanese. In *Second Language Speech*, ed. A. James and J. Leather. Mouton de Gruyter.
Yamada, R. A., and Tohkura, Y. 1990. Perception and production of syllable-initial English /r/ and /l/ by native speakers of Japanese. *Proceedings of the 1990 International Conference on Spoken Language Processing*, 757–60.
Yamada, R. A., and Tohkura, Y. 1991. Perception of American English /r/ and /l/ by native speakers of Japanese. In *Speech Perception, Production and Linguistic Structure*, ed. Y. Tohkura, E. Vatikiotis-Bateson, and Y. Sagisaka, editors. Ohmsha / Amsterdam: IOS Press.

Chapter • 11

Speech Perception, Language Acquisition, and Linguistics *Some Mutual Implications*

Henning Wode

Although it would be unfair to suggest that speech perception researchers have, in the past, paid little attention to the development of linguistics, the reverse, unfortunately, is quite accurate, as a cursory glance at introductory textbooks on linguistics will show. A fact common to most of these books is that speech perception is not treated. Worse, the phonological descriptions of individual languages and current phonological theories tend to be based on production phonology without, in general, acknowledging this fact. The purpose of this paper is to explore some of the results of speech perception research and to develop some implications for linguistics and vice versa.

Two perspectives are important for this chapter. One is a developmental one. That is, I look at things from the point of view of the learnability of languages and linguistic structures, respectively. The second perspective is a functional one. The evidence from speech perception research is reviewed from the point of view of the functions that speech perception has to meet with respect to the nature and overall functioning of natural human languages. In particular, I emphasize the importance of two kinds of abilities in speech perception; namely, to make gradual judgments as well as categorical ones. This focus leads to a number of issues including the relationship between the two abilities; the kinds of categorical distinctions an individual can develop; perceptuomotor links before the onset of speech; functional biological substrates for speech perception; the impact of perception on the evolution of speech; and the need for speech perception research to go beyond first- and second-language (L1 and L2) acquisition