音声研究　第15巻第 3 号
2011（平成23）年 12月
5–13頁

Journal of the Phonetic Society
of Japan, Vol. 15 No. 3
December 2011, pp. 5–13

研究ノート

Effects of L2 Experience on Perception of English /r/ and /l/ by Native Japanese Speakers

Katsura AOYAMA* and James E. FLEGE**

第二言語の使用経験が日本語話者の /r/ と /l/ の知覚に及ぼす影響

要旨：本研究ではアメリカに在住する日本語話者 50 名を対象に英語の /r/ と /l/ の知覚同化と弁別能力を検証した。実験 1 では英語の /r/ と /l/ の音が日本語のどの音の範疇に知覚同化されるか，また英語の /r/ と /l/ と日本語のラの音との類似性を，実験 2 では被験者の英語の /r/ と /l/ の弁別能力を検証した。実験 1 の結果から，英語の /r/ と /l/ は共に日本語のラの音に知覚同化されるが，日本語の音と英語の音との類似性を示す値と，アメリカ在住期間の長さの間に負の相関関係があることが示された。しかし，実験 2 の結果では，アメリカ在住期間の長さと英語の /r/ と /l/ の弁別能力の間に統計的に有意な相関関係は見られなかった。これらの結果から，アメリカでの在住期間が長い程，日本語話者が英語の /r/ と /l/ と，日本語のラの音声的な違いに気付いている可能性があるが，それが /r/ と /l/ の弁別能力にすぐ結びつくものではないということが示唆された。

Key words: English, Japanese, second language acquisition, speech perception, discrimination, length of residence

1. Introduction

Theories dealing with cross-language and second language perception have emphasized the importance of the perceived relation between phonetic segments in the listener's native language (L1) and those in a second or foreign (L2) language (Flege 1995, Best and Tyler 2007). More specifically, it is proposed that L2 phonetic segments are perceptually processed as if they were instances of an L1 category (Best 1995, Flege 1995, Best and Tyler 2007). The perceived relationship between the L2 phonetic segments and L1 phonological categories has often been assessed via two techniques; participants are first asked to categorize tokens of L2 segments in terms of L1 categories, and then to rate the tokens for similarity to the selected L1 category (e.g., Guion *et al.* 2000, Harnsberger 2000, Iverson *et al.* 2003).

It is well known that Japanese speakers have difficulty in perceiving English /r/ and /l/ (Goto 1971, Miyawaki *et al.* 1975, Takagi and Mann 1995, Yamada 1995, McCandliss *et al.* 2002, Aoyama *et al.* 2004). English /r/ (phonetically [ɹ], a central approximant) and English /l/ (phonetically a lateral approximant [l] or as velarized [ɬ], depending on context; Ladefoged 2006) are typically identified as instances of a single category, the so-called Japanese "r" (Takagi 1993, Komaki *et al.* 1999, Guion *et al.* 2000). The Japanese /r/ is often phonetically realized as an apico-alveolar tap [ɾ] (Price 1981, Vance 1987) but many allophones exist for this phoneme (Vance 1987). Guion *et al.* (2000) showed that Japanese speakers identified both English /r/ and /l/ tokens as poor examples of the Japanese /r/ category. In Iverson *et al.* (2003), Japanese speakers mostly identified synthetic tokens of both English /r/ and /l/ stimuli as instances of Japanese /r/.

Even though English /r/ and /l/ are typically identified as Japanese /r/, they may differ somewhat in the degree of similarity to Japanese /r/. In Takagi (1993), Japanese speakers gave lower ratings to /r/ than to /l/ when these sounds occurred in syllable-initial, initial cluster, and intervocalic positions. Komaki *et al.* (1999) also reported lower similarity ratings for English /r/ than /l/ with respect to Japanese /r/. In other words, English /r/ may be phonetically less similar to Japanese /r/ than English /l/ is.

An important question for L2 learning is whether perceived similarity will change as a function of experience. The Speech Learning Model (SLM, Flege 1995) focuses on L2 learning which occurs through experience

* Associate Professor, Department of Speech, Language & Hearing Sciences, Texas Tech University Health Sciences Center（テキサス工科大学健康科学センター，言語及び聴覚科学学部准教授）
** Professor Emeritus, University of Alabama at Birmingham（アラバマ大学バーミングハム校名誉教授）

with the L2 sound system. A key hypothesis of the SLM is that, although the phonetic elements of both the L1 and L2 are perceptually related to one another in a common phonological space, a new phonetic category can eventually be established for an L2 category if learners discern a sufficient amount of phonetic differences between the L2 category and its closest counterpart in the L1. This implies that category formation may, in some instances, depend on changes in perceived similarity that occur as L2 learners gain experience with particular L2 categories.

Some previous studies indicate that L2 segmental perception changes as a function of L2 experience (Yamada 1995, Flege *et al.* 1996 among others). Japanese speakers, for example, show an improvement in perception of English /r/ and /l/ as the result of formal phonetic training (e.g., Bradlow *et al.* 1997, 1999, Hazan *et al.* 2005, Iverson *et al.* 2005). Greater experience in using English in the real world may also produce similar effects. MacKain *et al.* (1981) found that Japanese adults who were relatively experienced in English identified a synthetic /r/-/l/ continuum more accurately than less experienced Japanese adults did. Flege *et al.* (1996) studied the identification of English /r/ and /l/ by Japanese adults differing greatly in length of residence (LOR) in the United States (U.S.). Japanese participants in the long LOR group (mean = 21 years) identified tokens of English /r/ and /l/ more accurately than did those in the short LOR group (mean = 2 years), although both groups obtained lower scores than native English speakers did.

The aims of this study were to examine the relationship of L2 experience with (a) perceived similarity of English /r/ and /l/ to Japanese /r/, and (b) native Japanese speakers' discrimination of English /r/ and /l/. The SLM hypothesizes that L2 learners will gradually discern phonetic differences between certain L1 and L2 categories, as they become more experienced in the L2 (Flege 1995). This view implies that perceptual similarity may change with increasing experience in the L2.

2. Experiment 1

Experiment 1 examined the perceptual similarity of English /r/ and /l/ with respect to Japanese /r/. Its aim was to determine whether perceptual similarity changes as a function of LOR. The stimuli consisted of naturally produced syllable-initial tokens of English /r/, /l/, and /w/. These stimuli were presented, along with syllable-initial Japanese /r/ and /w/ tokens, to Japanese participants differing in LOR in the U.S. English /w/ and Japanese /w/ were included because Best and Strange (1992) suggested that Japanese speakers may assimilate English /r/ and /l/ to Japanese /w/ and Yamada and Tohkura (1992) reported that Japanese speakers identified some of the synthetic /r/-/l/ continuum as /w/. The participants first rated English /r/, /l/, and /w/ for the degree of perceived similarity to Japanese /r/, and then classified the same stimuli as instances of Japanese consonant categories.

2.1 Method

2.1.1 Participants

A total of 53 Japanese adults were recruited in Birmingham, Alabama through advertisements in a university newspaper. All passed a pure-tone hearing screening at octave frequencies between 250 and 4000 Hz at 20 dB HL in both ears, and responded to a language background questionnaire prior to participating. No participant reported using a language other than English or Japanese.

The Japanese participants in this study were all born and raised in Japan. Many of them were students (n = 17) or research scientists (n = 11) at the University of Alabama at Birmingham. Other participants included a physician, workers at a Japanese-owned manufacturing facility, and spouses of researchers or other professionals.

To reduce heterogeneity in terms of age of first exposure to English (see Yamada 1995 for age effects on perception of /r/ and /l/), three participants who had spent some time in the U.S. prior to the age of 18 years were excluded from the study. The remaining 50 participants consisted of 14 men and 36 women, ranging in age from 21.2 to 48.3 years (mean = 33.2 years). Their age of arrival in the U.S. ranged from 19 to 39 years (mean = 29.4 years) and LOR ranged from 0.1 to 24.6 years (mean = 3.7 years). Among these 50 participants, majority (n = 40) of the participants' LOR was shorter than five years. We acknowledge that it would have been desirable to have a more equal distribution in LOR among participants. We believe that this pattern of LOR distribution among the participants is a reflection of a relatively small number of Japanese speakers in the Birmingham area.

2.1.2 Stimuli

The stimulus set consisted of five tokens each of English /r/, English /l/, English /w/, Japanese /r/, and Japanese /w/. The English stimuli consisted of consonant-vowel syllables produced by an adult male native speaker of American English. This speaker

produced seven syllables (/lɑ/, /rɑ/, /wɑ/, /bɑ/, /sɑ/, /θɑ/, /vɑ/)[1] multiple times in a carrier phrase (*Then I saw __ there*). The test sentences were recorded on a DAT recorder, then down-sampled to 22.05 kHz (16-bit amplitude resolution). The Japanese stimuli were produced by an adult male native speaker of Japanese. He produced /ra/ and /wa/ 18 times each in a carrier phrase similar to the one used in English (*Korewa __ desu*, "This is __"). There was no pause between the target syllable and the carrier phrase (e.g., English /rɑ/ and *saw* and *there*; Japanese /ra/ and *korewa* and *desu*). For both English and Japanese stimuli, five tokens of each syllable were edited out of the carrier phrase. They were normalized for peak intensity (50% of full scale).

In order to understand the acoustic nature of the English /rɑ/ and /lɑ/ stimuli used in this experiment, the following acoustic parameters were measured by using Pitchworks acoustic analysis software: consonant duration (in ms.), transition duration (in ms.), and second and third formant frequencies of the consonant (F2 and F3, in Hz). These parameters are reported as important cues in distinguishing English /r/ and /l/ (Polka and Strange 1985, Yamada and Tohkura 1992, Iverson *et al.* 2005). In addition, the overall stimuli duration (in ms.), vowel duration (in ms.), and F1 and F2 (in Hz) of the vowel were also measured. Segmentation criteria were based on both waveform and spectrogram cues as described by Kent and Read (2002) and Iverson *et al.* (2005). First, a steady state was identified for the vowel (/ɑ/) and the consonant (/r/ or /l/) using manual inspection of the spectrograms and automatic formant tracker in Pitchworks. Then the duration between the consonant steady state and vowel steady state was measured as transition duration. The formant frequencies were averaged across the steady state of the consonant or the vowel.

The acoustic characteristics of the English /r/ and /l/ stimuli are summarized in Table 1. As expected from the previous studies (Polka and Strange 1985, Yamada and Tohkura 1992, Iverson *et al.* 2005), F3 values were lower for English /r/ (mean = 1546 Hz) than for /l/ (mean = 2821 Hz). F2 did not appear to differ much between the two consonants (/r/ mean = 966 Hz, /l/ mean = 913 Hz). The average transition duration was longer for /r/ than for /l/ (means = 42 vs. 25 ms.), and the consonant steady-state duration was longer for /l/ than for /r/ (means = 98 vs. 69 ms.). Overall stimuli duration, vowel duration, and vowel F1 and F2 values were comparable between the English /r/ and /l/ stimuli.

Table 1 Acoustic characteristics (mean, s.d., range) of the English /r/ and /l/ stimuli in Experiments 1 and 2.

	English /r/	English /l/
Overall stimuli duration (ms.)	345 (32) 297–386	367 (54) 295–445
Consonant duration (ms.)	69 (16) 40–83	98 (32) 69–138
Transition duration (ms.)	42 (13) 33–64	25 (4) 21–30
F2 of the consonant (Hz)	966 (50) 903–1034	913 (43) 858–963
F3 of the consonant (Hz)	1546 (95) 1453–1704	2821 (22) 2794–2846
Vowel duration (ms.)	232 (22) 191–254	243 (28) 198–283
F1 of the vowel (Hz)	760 (5) 754–767	792 (15) 768–806
F2 of the vowel (Hz)	1154 (17) 1137–1185	1186 (22) 1157–1214

2.1.3 Procedure

The Japanese participants were tested individually in a sound booth at the University of Alabama at Birmingham. The Japanese participants first rated the consonant in each stimulus in terms of its degree of perceived similarity to Japanese /r/ using an equal-appearing interval scale that ranged from 1 (*not similar at all*) to 7 (*very similar*). The participants were told to ignore variations in the vowel insofar as possible, and to focus on the initial consonant in each stimulus. The 25 stimuli were randomly presented six times each. The interval between each response and the next stimulus was fixed at 1.0 s. The median of the final five ratings of each stimulus given by each participant for each consonant token was determined. Then an average was computed for the five tokens of each target consonant. This yielded a total of 240 mean ratings (each based on 25 judgments) for analysis.

Classification data were obtained after the rating task was completed. The same 25 stimuli were randomly presented three times each for forced-choice identification. Four response alternatives were ワ /wa/, ラ /ra/, ﾘｬ /rʲa/, or ヴァ /va/[2]. These were the response alternatives used by native Japanese speakers to identify English /r/ and /l/ in Guion *et al.* (2000)[3].

All instructions for this experiment, both written and

oral, were given in Japanese because bilingual speakers' perception may vary according to which language is used in the context (Elman *et al.* 1977). The participants were told to identify each stimulus using one of the four response alternatives presented on a computer screen in Japanese Katakana orthography. They were told that it was not necessary to use all response alternatives, or to use the alternatives with equal frequency. A response was required for each stimulus, and participants were told to guess in case of uncertainty. A stimulus could be replayed, but responses could not be changed once given. The interval between each response and the next trial was fixed at 1.0 s.

The final two classifications of each stimulus were retained for analysis, yielding 2500 identification judgments (50 participants × 5 consonants × 5 tokens × 2 replicate judgments). Correlations between similarity ratings of English /r/ and /l/ and the participants' LOR were examined.

2.2 Results

2.2.1 Classification data

The identification responses obtained for the English /lɑ/, /rɑ/, /wɑ/ and Japanese /ra/ and /wa/ stimuli have been tabulated in Table 2. The percentages for each consonant were based on 500 responses (50 participants × 5 stimulus tokens × 2 replicate judgments).

Participants nearly always (98.8%) classified the Japanese /ra/ tokens as Japanese /ra/. This verified that they understood the task, and provided a point of reference for the classifications of the English /r/ and /l/ tokens. The modal classification of both the English /rɑ/ and /lɑ/ tokens was Japanese /ra/ (ラ). The average frequency with which English /lɑ/ tokens were classified as Japanese /ra/ was 81.6%; the average frequency for the English /rɑ/ tokens was 77.0%. When the English liquids were not classified as the Japanese /ra/ (ラ), it was usually classified as ヴァ /va/ (means = 12.8% for /l/, 17.2% for /r/). The Japanese /wa/ tokens were most often classified as ワ /wa/ (mean = 80.2%). The English /wɑ/ tokens were also most often classified as ワ /wa/ (mean = 77.6%). The English /lɑ/ tokens were never classified as Japanese /wa/, and the English /rɑ/ tokens were very seldom classified as Japanese /wa/ (mean = 1.2%).

Table 2 The mean identification and similarity rating of English and Japanese consonants that were identified in terms of Japanese consonants in Experiment 1.

Stimulus	Response			
	ワ (wa)	ラ (ra)	リャ (rʲa)	ヴァ (va)
E /l/		**81.6** (4.8)	5.6	12.8
E /r/	1.2	**77.0** (4.4)	4.6	17.2
E /w/	**77.6**	2.4 (1.6)		19.6
J /r/		**98.8** (6.8)		
J /w/	**80.2**	2.0 (1.8)		17.4

Note: E, English; J, Japanese; boldface values indicate the modal response. The percentages for each consonant were based on 500 judgments (50 participants × 5 tokens × 2 judgments). Percentages less than 1% are not shown. The mean similarity ratings to Japanese /r/ are in parentheses underneath the mean identification percentages as Japanese ラ (ra).

2.2.2 Rating data

Table 2 also shows the mean similarity ratings. The Japanese /r/ tokens were judged to be very similar to the Japanese /r/ category (mean = 6.8 on the 7-point scale). The Japanese /w/ and English /w/ tokens received very low mean similarity ratings (mean = 1.8 and 1.6 respectively). The mean similarity ratings obtained for both English liquids were intermediate to those ratings obtained for the Japanese /r/ tokens and the /w/ tokens. The mean similarity ratings for the English /r/ tokens averaged 4.4, and the ratings for the English /l/ tokens averaged 4.8.

The effect of LOR was examined by correlations. The negative correlation between LOR and the per-

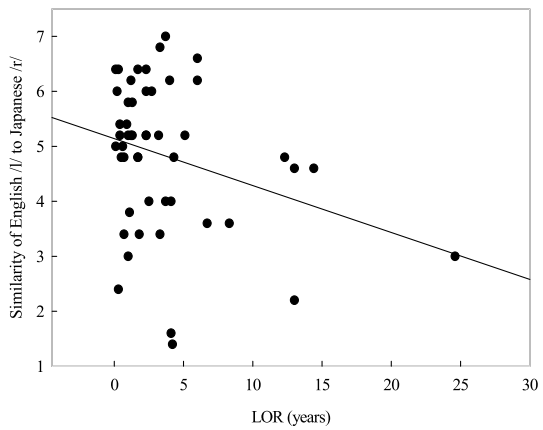

Fig. 1 Correlation between LOR and similarity of English /l/ to Japanese /r/ (N = 50)

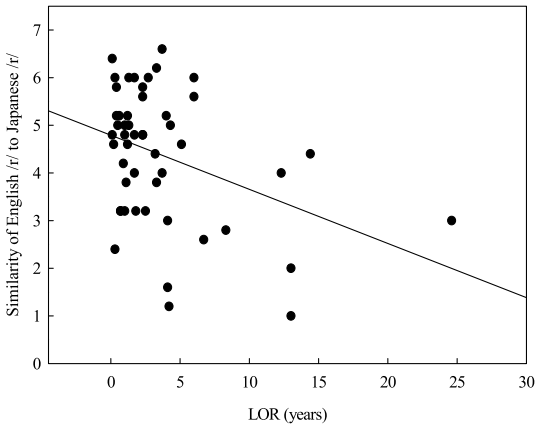

Fig. 2 Correlation between LOR and similarity of English /r/ to Japanese /r/ (N = 50)

ceived similarity ratings of English /l/ was statistically significant ($r(48) = -0.29$, $p = 0.039$) (see Figure 1). The negative correlation between LOR and the perceived similarity ratings of English /r/ was also statistically significant ($r(48) = -0.38$, $p = 0.007$) (see Figure 2). These significant negative correlations indicate that the Japanese speakers rated English /r/ and /l/ as less similar to Japanese /r/ as their LOR increased[4].

In summary, the results from Experiment 1 confirm the two-to-one cross-language perceptual mapping pattern reported in previous research (Takagi 1993, Komaki *et al.* 1999, Guion *et al.* 2000). Correlation analysis indicated that the similarity ratings of English /r/ and /l/ decreased as the participants' LOR increased. This suggests that, the longer the Japanese speakers lived in the U.S., the more likely that they noticed the phonetic differences between English /r/ and /l/ and Japanese /r/.

3. Experiment 2

The aim of Experiment 2 was to determine if the Japanese participants' discrimination accuracy of English /r/-/l/ would increase as LOR increased. The participants from Experiment 1 were tested on their discrimination of English /r/-/l/, /r/-/w/, /s/-/θ/, /b/-/v/, all of which have been shown to present perceptual difficulty for Japanese speakers of English (Yoshida and Hirasaka 1983, Best and Strange 1992, Guion *et al.* 2000). The fifth contrast, /b/-/s/, served as a control. It was expected that the Japanese participants obtain high discrimination scores for the /b/-/s/ contrast because the Japanese phonemic inventory includes both /b/ and /s/ (Vance 1987), and because these two sounds differ in voice, places of articulation, and manners of articulation.

3.1 Method

3.1.1 Participants

The Japanese adults described earlier participated in this experiment shortly after completing Experiment 1.

3.1.2 Stimuli and Procedure

The English stimuli used in Experiment 1 were used in a categorial discrimination test that has been used previously in L2 research (Guion *et al.* 2000, Aoyama *et al.* 2004, 2008) because of its suitability as a test of category formation. Each trial consisted of three physically different stimuli separated by an inter-stimulus interval of 0.5 s. The five contrasts of interest (/b/-/s/, /b/-/v/, /r/-/l/, /r/-/w/, /s/-/θ/) were each tested by eight *change* and eight *no-change* trials. The participants' task was to select an odd item out if they judged one stimulus to be different from the other two, or to respond *No* if they did not (i.e., for the no-change trials). More specifically, the correct response to change trials was a button (*1*, *2*, or *3*) indicating the position of the odd item out, which occurred with near equal frequency in all three possible serial positions. The no-change trials consisted of three physically different tokens of a single category. The correct response to these trials was a fourth button marked *No* (for no odd item out).

The change trials tested the participants' ability to distinguish consonants drawn from different phonetic categories, whereas the no-change trials tested their ability to ignore audible but phonetically irrelevant within-category variation. No-change trails were included in this experiment to test category formation in L2, not simply the auditory discrimination between two sounds. In order to obtain a high score on this test, the participants needed to ignore within-category variations and also detect the difference between tokens that were drawn from separate categories.

The test was administered by the experimenter from Experiment 1, who used English with all participants in Experiment 2 because the discrimination of English consonants was being tested. The stimuli were presented via headphones. Before the experiment began, a presentation level was established by having each participant adjust the volume level using stimuli similar to those of the test. Once a comfortable level had been selected, it was fixed for the remainder of the experiment.

Next, the participants took part in a practice session with feedback, during which /wɑ/ and /sɑ/ stimuli were presented in triadic trials. The participants were required to respond correctly to at least nine of ten practice items before proceeding to the experiment proper. All participants were able to meet this criterion in fewer than four practice blocks. Feedback was not presented during the experiment. The trials testing all five contrasts were presented in a single randomized block. Ten extra trials at the beginning of the block were presented for practice, but were not analyzed.

The dependent variable examined in analyses were *A'* (*A-prime*) values calculated for each participant. *A'* is one of the measures used in signal detection theory (Snodgrass *et al.* 1985). The *A'* is based on the proportion of *hits* (correct selections of the odd item out in change trials) and *false alarms* (incorrect selections of the odd item out in no-change trial) obtained for each contrast. The formula provided by Snodgrass *et al.* (1985) was used. There was a maximum of eight hits and eight false alarms for each contrast. If the proportion of hits (H) equaled the proportion of false alarms (FA), *A'* was set to 0.5. If H exceeded FA, then

$A' = 0.5 + ((H - FA)*(1 + H - FA))/((4*H)*(1 - FA))$

If FA exceeded H, then

$A' = 0.5 - ((FA - H)*(1 + FA - H))/((4*FA)*(1 - H))$

A score of 1.0 was obtained when responses were correct for all 16 trials testing a contrast. A score of 0.5 indicated a theoretically defined chance level of response (see Snodgrass *et al.* 1985). Correlations between *A'* scores and the participants' LOR were examined.

3.2 Results

The average *A'* score for the control contrast, /b/-/s/, was 0.98. This demonstrated that the Japanese participants understood the instructions. The average *A'* scores for /r/-/w/, /b/-/v/, and /s/-/θ/ were 0.86, 0.83, and 0.80, respectively. The lowest average *A'* score was obtained for /r/-/l/ (0.60). For a comparison, the average *A'* scores obtained by adult native English speakers (n = 16) were /b/-/s/ (0.97), /r/-/w/ (0.96), /b/-/v/ (0.91), and /s/-/θ/ (0.94), /r/-/l/ (0.95) (Aoyama *et al.*, 2004, 2008).

The effect of LOR was examined by correlations. The correlations between LOR and *A'* scores for /b/-/s/, /b/-/v/, /r/-/l/, and /r/-/w/ were non-significant ($r(48) = -0.03$ to 0.15, $p > 0.1$) (see Figure 3 for correlation between LOR and /r/-/l/ scores). The correlation between LOR and *A'* scores for /s/-/θ/ was statistically significant ($r(48 = 0.30$, $p = 0.036$). The significant positive correlation between LOR and *A'* scores for /s/-/θ/ indicate that the Japanese speakers with longer LOR obtained higher discrimination scores for this contrast. The lack of significant correlation for /r/-/l/, /b/-/s/, /b-/v/, and /r/-/w/ indicates that the Japanese speakers with differing LOR obtained comparable *A'* scores for these contrasts [5].

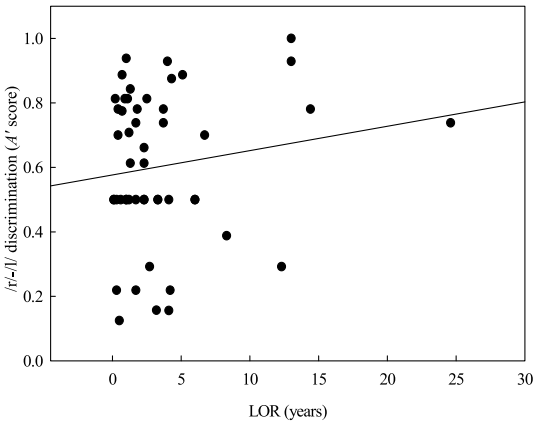

Fig. 3 Correlation between LOR and /r/-/l/ discrimination scores (*A'*) (N = 50)

The results of Experiment 2 agreed with the findings of Takagi and Mann (1995) in suggesting that Japanese speakers with a relatively long LOR continue to have difficulty with /r/-/l/ discrimination. The results indicate that the participants with longer LOR discriminated /s/ from /θ/ more accurately than those with shorter LOR. This suggests that discrimination ability of at least some L2 segmental contrasts improves with increasing LOR.

4. General discussion

The primary aim of this study was to investigate the effects of L2 experience on Japanese speakers' perception of English /r/ and /l/. Specifically, this study examined the effects of LOR in the U.S. experience on perceived similarity of English /r/ and /l/ to Japanese /r/, and on discrimination of English /r/ and /l/ by native Japanese speakers. The results of Experiment 1 indicated that, overall, Japanese speakers classified both English /r/ and /l/ as the Japanese /r/ category, and that the mean similarity ratings obtained for both English liquids were in between those obtained for the Japanese /r/ and /w/ tokens. It was found that LOR and similarity ratings of English /r/ and /l/ to Japanese /r/ were both negatively correlated, suggesting that more experienced

Japanese speakers may be able to discern phonetic differences between English liquids and Japanese /r/ more accurately than less experienced Japanese speakers. The results of Experiment 2 showed that the participants' LOR did not correlate with their discrimination scores of English /r/-/l/. Taken together, the results suggest that Japanese speakers with more L2 experience may be more sensitive to phonetic differences between English liquids and Japanese /r/ than those with less experience, but that the participants with more L2 experience did not discriminate English /r/-/l/ more accurately than those with less L2 experience.

The question is why discrimination scores of /r/-/l/ did not increase with longer LOR. Is it because the /r/-/l/ contrast is so difficult for Japanese speakers that their discrimination ability does not improve by just living in the U.S.? As mentioned earlier, one problem in this study is that the participants' LOR values were not evenly distributed. As can be seen in Figure 3, many of the participants had lived in the U.S. for less than five years. The four participants who had the longest LOR scored relatively high (>0.74) on the discrimination of /r/-/l/, and those with relatively long LOR had *A'* scores higher than 0.5. It was also pointed out some of the Japanese speakers might have paid attention to acoustic cues that were not necessarily important for distinguishing /r/-/l/. For instance, Iverson *et al.* (2005) reported that transition and closure durations affected Japanese speakers' identification of /r/ and /l/. Attention to other acoustic cues could lead Japanese speakers to have higher false alarm rates, which would in turn lower their *A'* scores.

In addition, there was a wide variety of discrimination scores among the participants who had lived in the U.S. for a relatively short time (see Figure 3). Among the participants whose LOR was shorter than five years, the discrimination scores ranged from 0.13 to 0.94. In other words, some of the Japanese speakers discriminated English /r/-/l/ fairly accurately despite living in the U.S. for a short period of time. Clearly, there is something other than LOR that affected Japanese speakers' ability to discriminate English /r/ and /l/. Qualitative research of "good" and "poor" discriminators among these participants with a short LOR may reveal what other factors affect L2 perception.

It is also possible that the effect of LOR on L2 perception may be observed more easily in other L2 contrasts, such as English /s/-/θ/ for Japanese speakers. The results of Experiment 2 showed that LOR and discrimination scores were positively correlated for the /s/-/θ/ contrast, although LOR and discrimination scores did not correlate significantly for the other four contrasts. The results of Experiment 1 also indicated that Japanese speakers with longer LOR are more aware of the phonetic differences between English liquids and Japanese /r/. These results suggest that LOR does affect the perception of L2 sounds, but discrimination of English /r/-/l/ might be particularly difficult for native Japanese speakers (see Goto 1971, Takagi and Mann 1995; also Larson-Hall 2006 for the production of /r/ and /l/).

Another possibility is that LOR alone may not have been a good indicator of L2 experience. Flege and Liu (2001) found LOR effects among students but not among non-students in Chinese L2 learners of English. The students with longer LOR (mean = 7.3 years) had higher scores than the students with shorter LOR (mean = 2.5 years) on consonant identification, grammaticality judgments, and listening comprehension. In contrast, the difference between the two LOR groups (means = 1.7 vs. 6.6 years) was not found among non-students. Flege and Liu (2001) suggested that the degree of L2 learning depended on the quality of input, and that the students had more opportunities to interact with native speakers of English than the non-students. In addition, some previous studies demonstrated that short-term laboratory training can yield improvements in Japanese speakers' perception of /r/ and /l/ (e.g., Strange and Dittmann 1984, Bradlow *et al.* 1997, 1999, Iverson *et al.* 2005), and improvements in production after a perceptual training (Bradlow *et al.* 1997, 1999). Thus, the nature of L2 experience (e.g., native speaker input, laboratory training in speech perception) may need to be examined in addition to LOR as an index of L2 experience in future research.

The acoustic characteristics of the /r/ and /l/ stimuli differed slightly in comparison to those used in Yamada and Tohkura (1992) and Iverson *et al.* (2005) in that English /r/ and /l/ had similar F2 frequencies (966 Hz for /r/ and 913 Hz for /l/, see Table 1). In Yamada and Tohkura (1992), the F2 values were manipulated from 920 Hz for "typical /r/" to 1280 Hz for "typical /l/" and the average F2 frequency for /l/ was higher than 1500 Hz in Iverson *et al.* (2005). However, the F3 frequencies in the tokens used in this study were comparable (1546 Hz for /r/ and 2821 Hz for /l/) to these two previous studies. It is unlikely that the characteristics of the /r/ and /l/ stimuli had an effect on the Japanese speakers' discrimination scores, especially because the native English speakers in Aoyama *et al.* (2004) had high discrimination scores (mean = 0.95) for the /r/-/l/ contrast using the exact same stimuli and procedure.

The present research is one of the first attempts to evaluate the effect of L2 experience on both perceived similarity and discrimination of L2 sounds (see Levy 2009 for perception of French vowels and the effect of experience). The results of this study suggested that additional LOR in the U.S. had relatively little effects on Japanese speakers' discrimination of English /r/ and /l/, although Japanese speakers with longer LOR were more sensitive to the phonetic differences between English /r/ and Japanese /r/. It was also found that the Japanese speakers with longer LOR were able to discriminate another L2 contrast, /s/-/θ/, more accurately than those with shorter LOR. Future research is needed to determine when and how changes in perceived similarity occur for different L2 contrasts, and when these changes begin to impact the discrimination ability of phonetic categories.

Acknowledgements

This study was supported by NIH grant DC00257. The funding source had no involvement in study design, data collection, analysis, and interpretation of the data. The authors thank J. Dembowski, S. Bayona, E. Cundari, K. Bradshaw, and K. Rich for their assistance, and the editor and two anonymous reviewers for their suggestions. All procedures were approved by the Institutional Review Board of the University of Alabama at Birmingham, and informed consent was obtained from every participant.

Notes

1) Consonants other than /l/, /r/ and /w/ were also recorded because the recordings were for stimuli in the consonant discrimination test (Experiment 2) as well. These stimuli were also used in a consonant discrimination test in Aoyama *et al.* (2004).
2) It was suggested that ダ (/da/) should have been included because native Japanese-speaking children and adults often confuse Japanese /r/ with Japanese /d/ (see Otsuka 2005). For Experiment 1, it is possible that Japanese speakers would have identified Japanese /ra/ as ダ, had the choice been offered. Unfortunately, it is no longer possible to retest the participants in the current study with choices that include ダ. It should be noted that no Japanese participants mentioned that they wished another choice, such as ダ, was available, and that Japanese /ra/ was identified as intended (ラ) 98.8% of time across participants.
3) We did not offer ウラ /ura/ as a response alternative, because the primary aim of this experiment was to learn which consonant category the Japanese speakers use to identify each stimulus.
4) As described in 2.1.1, the majority of the participants lived in the U.S. for less than five years. It was pointed out that the uneven distribution of the participants' LOR potentially affected the results and interpretation of the data. Indeed, when the participant with the longest LOR was taken out, the correlation between LOR and the similarity of /l/ was not significant ($r(47) = -0.223$, $p = .12$). The correlation between LOR and the similarity of /r/ was significant ($r(47) = -.381$, $p = .007$).
5) The distribution of the participants' LOR may have affected the results here as well. This problem will be discussed in section 4.

References

Aoyama, K., J. E. Flege, S. G. Guion, R. Akahane-Yamada and T. Yamada (2004) "Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/." *Journal of Phonetics* 32, 233–250.

Aoyama, K., S. G. Guion, J. E. Flege, T. Yamada and R. Akahane-Yamada (2008) "The first years in an L2-speaking environment: A comparison of Japanese children and adults learning American English." *International Review of Applied Linguistics in Language Teaching* 46, 61–90.

Best, C. T. (1995) "A direct realist view of cross-language speech perception." In W. Strange (ed.) *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, 171–204. Timonium, MD: York Press.

Best, C. T. and W. Strange (1992) "Effects of phonological and phonetic factors on cross-language perception of approximants." *Journal of Phonetics* 20, 305–330.

Best, C. T. and M. D. Tyler (2007) "Nonnative and second-language speech perception." In O.-S. Bohn and M. J. Munro (eds.) *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*, 13–34. Amsterdam: John Benjamins.

Bradlow, A., D. B. Pisoni, R. Akahane-Yamada and Y. Tohkura (1997) "Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production." *Journal of the Acoustical Society of America* 101, 2299–2310.

Bradlow, A. R., R. Akahane-Yamada, D. B. Pisoni and Y. Tohkura (1999) "Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production." *Perception and Psychophysics* 61, 977–985.

Elman, J. L., R. L. Diehl and S. E. Buchwald (1977) "Perceptual switching in bilinguals." *Journal of the Acoustical Society of America* 62, 971–974.

Flege, J. E. (1995) "Second language speech learning:

Theory, findings, and problems." In W. Strange (ed.) *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, 233–277. Timonium, MD: York Press.

Flege, J. E., N. Takagi and V. Mann (1996) "Lexical familiarity and English-language experience affect Japanese adults' perception of /ɹ/ and /l/," *Journal of the Acoustical Society of America* 99, 1161–1173.

Flege, J. E. and S. Liu (2001) "The effect of experience on adults' acquisition of a second language." *Studies in Second Language Acquisition* 23, 527–552.

Goto, H. (1971) "Auditory perception by normal Japanese adults of the sounds "L" and "R"." *Neuropsychologia* 9, 317–323.

Guion, S. G., J. E. Flege, R. Akahane-Yamada and J. C. Pruitt (2000) "An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants." *Journal of the Acoustical Society of America* 107, 2711–2724.

Harnsberger, J. D. (2000) "A cross-language study of the identification of non-native nasal consonants varying in place of articulation," *The Journal of the Acoustical Society of America* 108, 764–783.

Hazan, V., A. Sennema, M. Iba and A. Faulkner (2005) "Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English." *Speech Communication* 47, 360–378.

Iverson, P., P. Kuhl, R. Akahane-Yamada, E. Diesch, Y. Tohkura, A. Kettermann and C. Siebert (2003) "A perceptual interference account of acquisition difficulties for non-native phonemes." *Cognition* 87, B47–B57.

Iverson, P., V. Hazan and K. Bannister (2005) "Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults." *Journal of the Acoustical Society of America* 118, 3267–3278.

Kent, R. D. and C. Read (2002) *Acoustic Analysis of Speech* (2nd ed.). Albany, NY: Singular Thomson Learning.

Komaki, R., R. Akahane-Yamada and Y. Choi (1999) "Effects of native language on the perception of American English /r/ and /l/: Cross-language comparison between Korean and Japanese." *Technical Report of IEICE, SP 99-45* (pp. 39–46). Tokyo: The Institute of Electronics, Information and Communication Engineers.

Ladefoged, P. (2006) *A Course in Phonetics* (5th ed.). Boston: Thomson Wadsworth.

Larson-Hall, J. (2006) "What does more time buy you? Another look at the effects of long-term residence on production accuracy of English /r/ and /l/ by Japanese speakers." *Language and Speech* 49, 521–548.

Levy, E. S. (2009) "Language experience and consonantal context effects on perceptual assimilation of French vowels by American-English learners of French." *Journal of the Acoustical Society of America* 125, 1138–1152.

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

McCandliss, B., J. Fiez, A. Protopapas, M. Conway and J. McClelland (2002) "Success and failure in teaching the [r]-[l] contrast to Japanese adults: Predictions of a Hebbian model of plasticity and stabilization in spoken language perception." *Cognitive, Affective and Behavioral Neuroscience* 2, 89–108.

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

Otsuka, N. (2005) "Articulation development and phonological perception." *Nihon Daigaku Daigakuin Soogoo Shakai Joohoo Kenkyuuka Kiyoo* 6, 150–160.

Polka, L. and W. Strange (1985) "Perceptual equivalence of acoustic cues that differentiate /r/ and /l/." *Journal of the Acoustical Society of America* 78, 1187–1197.

Price, P. (1981) "A Cross-Linguistic Study of Flaps in Japanese and in American English." Ph.D. dissertation, University of Pennsylvania.

Snodgrass, J. G., G. Levy-Berger and M. Haydon (1985) *Human Experimental Psychology*. New York: Oxford University Press.

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

Takagi, N. (1993) "Perception of American English /r/ and /l/ by Adult Japanese Learners of English: A Unified View." Ph.D. dissertation, University of California at Irvine.

Takagi, N. and V. A. Mann (1995) "The limits of extended naturalistic exposure on the perceptual mastery of English /r/ and /l/ by adult Japanese learners of English." *Applied Psycholinguistics* 16, 379–405.

Vance, T. (1987) *An Introduction to Japanese Phonology*. New York: State University of New York Press.

Yamada, R. A. (1995) "Age and acquisition of second language speech sounds perception of American English /ɹ/ and /l/ by Native speakers of Japanese." In W. Strange (ed.) *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, 305–320. Baltimore: York Press.

Yamada, R. A. and Y. Tohkura (1992) "The effects of experimental variables on the perception of American English /r/ and /l/ by Japanese listeners." *Perception and Psychophysics* 52, 376–392.

Yoshida, K. and F. Hirasaka (1983) "The lexicon in speech perception." *Sophia Linguistica* 11, 105–116.

(Received Mar. 2, 2010, Accepted Dec. 7, 2011)