

EFFECTS OF NATIVE LANGUAGE ON THE PERCEPTION OF AMERICAN ENGLISH /r/ AND /l/: A COMPARISON BETWEEN KOREAN AND JAPANESE

Ryo Komaki* and Youngon Choi[†]

*ATR Human Information Processing Research Laboratories, Kyoto, Japan

[†]Duke University, Durham, NC, USA

ABSTRACT

This paper attempted to clarify the nature of the difficulties that second language (L2) learners have. Three experiments were conducted to examine the effects of the first language (L1) on the perception of L2 speech contrasts. Specifically, perception of the American English (AE) /r/-/l/ contrast by Korean speakers and Japanese speakers was compared, and the following results were obtained. First, Korean speakers' overall perception ability was significantly higher than Japanese speakers'. However, Korean speakers' identification was biased toward /l/ responses while Japanese speakers' was not. Second, language-specific patterns were found in the identification task, i.e., syllabic positions difficult for Korean speakers were different from Japanese speakers. The pattern for Korean speakers was found to be explained by the perceptual assimilation of AE /r/-/l/ to Korean phonetic categories. The results are discussed in the context of current theories of cross-language speech perception.

1. INTRODUCTION

Current cross-language studies have focused on first-language (L1) interference in the perception of foreign speech segments. For example, Ingram and Park [1] reported that speakers of Korean and Japanese both have difficulties perceiving the Australian English /r/-/l/ contrast, but not in the same way. They compared the identification ability between the two language groups using minimal pair English words contrasting /r/ and /l/ in three syllabic positions (initial singleton, initial cluster, and inter vocalic position) and found a language-specific pattern in their difficulty.

Lively et al. [2] examined the syllable position effect in Japanese speakers' perception of American English (AE) /r/ and /l/. They used five positions: the three positions that Ingram and Park used plus final cluster and final singleton. They found that /r/ and /l/ in the two final positions were easier to identify for Japanese listeners than other three positions. This suggests that the Korean-Japanese comparison should be re-examined using all five phonetic positions, in order to further analyzing the L1 effect on /r/-/l/ acquisition.

Recently, Best [3] hypothesized in her Perceptual Assimilation Model (PAM) that the assimilation patterns of non-native contrasts into native categories are predictive of difficulties that non-native listeners have. However, few studies have examined perceptual assimilation directly. Given the notion that L1 interferes with non-native speech perception, it is necessary to assess perceptual assimilation in order to understand difficulties that L2 learners have.

The present study aims to further understand the effect of L1 in L2 learning. Korean and Japanese adults' perception of the AE /r/-/l/ contrast was compared. In the first experiment, identification performance as a function of position was compared between the two language groups. To do so, we expanded Ingram and Park's

study by adding two more syllabic positions: final cluster and final singleton. In the second experiment, we compared the shape of the perceptual boundary between /r/ and /l/ in a categorical perception experiment. In the third experiment, perceptual assimilation of AE /r/ and /l/ to Korean phonetic categories was assessed.

2. EXPERIMENT 1

2.1. Method

2.1.1. Subjects. Twenty-two native speakers of Korean (11 males and 11 females) and 40 native speakers of Japanese (20 males and 20 females) participated in the experiment. Korean speakers were undergraduate students at Seoul National University and Ewha Womans University in Seoul, Korea, and were 22 years of age on average. Japanese speakers were undergraduate students at Doshisha University in Kyoto, Japan, and were 19 years of age on average. None had lived abroad, but had received English education in high school and in college for several years. None of the subjects reported any history of speech or hearing impairment.

2.1.2. Stimuli. The stimulus set consisted of 53 pairs of English words minimally contrasting in /r/ and /l/ in one of five phonetic environments: initial singleton (IP), initial consonant cluster (IC), intervocalic (IN), final consonant cluster (FC), and final singleton (FP). Each word was read by one of four speakers of American English (two males and two females). Talkers A (male) and B (female) read 15 minimal pairs, talker C (male) read 14 pairs, and talker D (female) read 9 pairs, for a total of 106 stimuli (53 pairs \times 2 per pair). The recordings were digitized at 16-bit resolution and 22.05 kHz sampling frequency, and each word was saved into a sound file.

2.1.3. Procedure. A two-alternative forced-choice (2AFC) identification task was used. Each trial began with a presentation on the computer display of one /r/-/l/ minimal pair in regular English orthography. Then, one member of the pair was presented auditorily over headphones. Subjects selected one of the two words by clicking the word on the display. Stimuli were presented in four talker blocks. A fixed order was used for sequencing the talkers; each block alternated between a male and a female. Within each talker block, each stimulus was presented once in a random order. There were 106 trials across the four blocks.

2.2. Results

Figure 1 shows the correct identification rate averaged over Korean subjects and Japanese subjects. A one-way ANOVA showed that Korean subjects showed significantly higher accuracy (78.4%) than Japanese subjects (63.9%) [$F(1, 60) = 32.8, p < 0.001$].

Effects of talker (talkers A, B, C and D), consonant (/r/ words

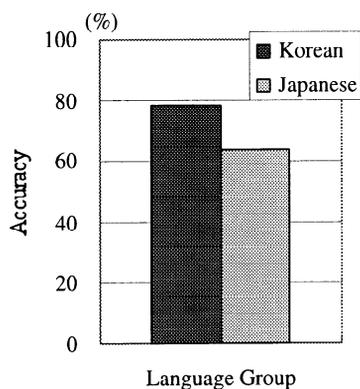


Figure 1: Correct identification rate averaged over Korean subjects and Japanese subjects.

vs. /l/ words), and phonetic environment (IP, IC, IN, FC and FP), along with their interaction with L1 (Korean vs. Japanese) were tested in three separate ANOVAs where the accuracy was the dependent variable.

2.2.1. Effect of talker. A two-way ANOVA with talker and L1 as independent variables showed a significant main effect of L1 [$F(1, 60) = 11.29, p < 0.005$]. The main effect of talker [$F(3, 180) = 2.60, N.S.$] and the interaction between L1 and talker [$F(3, 180) = 0.27, N.S.$] were not significant.

2.2.2. Effect of consonant. A two-way ANOVA with consonant and L1 as independent factors showed significant main effects of L1 [$F(1, 60) = 11.44, p < 0.005$] and consonant [$F(1, 60) = 8.30, p < 0.01$]. There was no significant interaction between L1 and consonant [$F(1, 60) = 2.95, N.S.$].

2.2.3. Effect of phonetic environment. Figure 2 shows the correct identification rate by phonetic environment averaged over Korean subjects and Japanese subjects. A two-way ANOVA with position and L1 as variables showed a significant main effect of L1 [$F(1, 60) = 12.45, p < 0.001$] and a significant position-by-L1 interaction [$F(4, 240) = 13.84, p < 0.001$]. The main effect of position was not significant [$F(4, 240) = 1.94, N.S.$].

3. EXPERIMENT 2

3.1. Method

3.1.1. Subjects. Subjects were the same as those in Experiment 1. All of them took part in Experiment 2 after they completed Experiment 1.

3.1.2. Stimuli. The stimuli were a synthetic /rait/-/lait/ continuum generated with the Klatt cascade formant synthesizer used in Yamada and Tohkura [4]. Each stimulus contained an initial steady portion of /r/-/l/ and a transition portion, followed by /ait/. Tracks of the first three formants (F1, F2 and F3) in the initial steady portion and transition portion were manipulated; they covaried in 17 steps from /r/-like characteristics (ST1) to /l/-like characteristics (ST17). Transition duration of F1 varied from 61 ms to 13 ms in 3 ms steps. F2 onset frequencies varied from 960 Hz to 1280 Hz in 20 Hz steps, and F3 onset frequencies varied from 1400 Hz to 3000 Hz in 100 Hz steps.

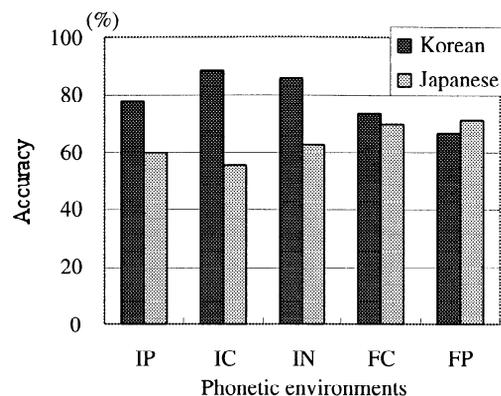


Figure 2: Correct identification rate as a function of phonetic environment.

3.1.3. Procedure: Identification task. A three-alternative forced-choice (/r/, /l/, and /w/) identification task was used. Each trial began with the presentation of three response buttons, each representing R, L, and W, on the computer display. Then, one of the 17 stimuli was presented auditorily over headphones. Subjects identified the initial consonant and responded by clicking one of the buttons. The experiment was a self-paced form in which subjects' responses were followed by presentation of the next stimulus. The seventeen stimuli were presented 10 times in a random order.

3.1.4. Procedure: Discrimination task. Thirteen four-step comparison pairs (ST1-ST5, ST2-ST6, ..., ST13-ST17) were chosen from the 17 stimuli. Each pair was arranged in triads in four possible ABX permutations, ABA, ABB, BAA, and BAB, resulting in 52 (13 A-B pairs \times 4 ABX forms) triads. Each triad was presented four times in a random order for a total of 208 trials. Each trial began with the presentation of response buttons ("A" and "B") on the computer display. Then, one of the 52 triads was presented auditorily over headphones with an inter-stimulus-interval (ISI) of one second. Subjects judged whether the third item (X) in a triad matched the first item (A) or the second item (B), by clicking "A" or "B". Subjects' responses were followed by presentation of the next triad.

3.2. Results and Discussion

Figure 3 shows the results from the identification test (top panels) and the discrimination test (bottom panels). Generally, Korean subjects showed better performance than Japanese.

In both tests, responses averaged over Japanese subjects were close to chance level, i.e., 33.3% in identification and 50% in discrimination. Several studies have reported that Japanese speakers do not perceive the /r/-/l/ continuum categorically. The present results replicated this finding.

In contrast, Korean subjects' identification was fairly consistent, but their response pattern was highly biased toward /l/ responses. The discrimination function showed a peak around ST7-ST11. Taken together, Korean subjects perceived the continuum somewhat categorically, but by categorizing them into two classes, non-/l/ vs. /l/.

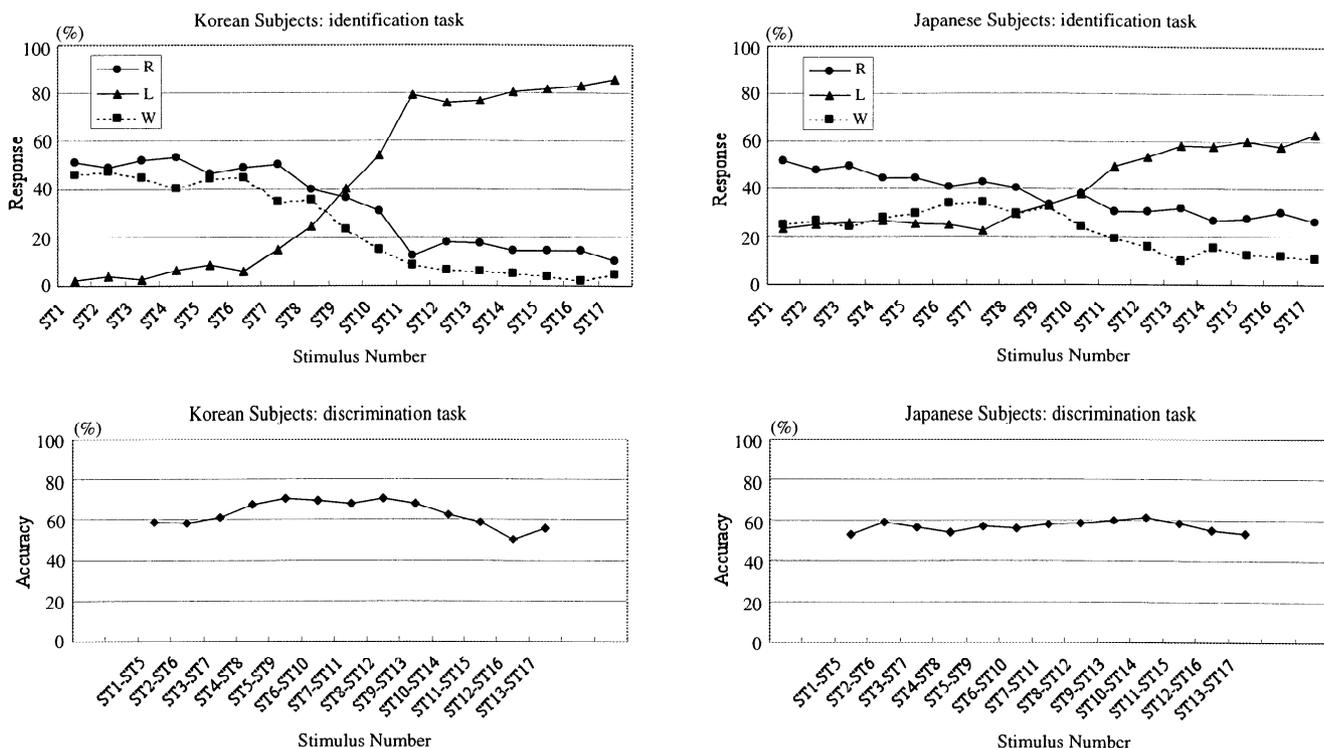


Figure 3: The results of Experiment 2. The top panels show responses for R, L, and W in the identification task. The bottom panels show percent correct responses in the discrimination task. The left and right panels show the results from Korean and Japanese speakers, respectively.

4. EXPERIMENT 3

Differences in the accuracy pattern by position between Korean and Japanese subjects demonstrated in Experiment 1 suggest that L1 interferes with L2 perception. In Experiment 3, we examined whether the above finding can be explained by PAM. To do so, we assessed the perceptual assimilation of AE /r/-/l/ to two possible Korean phonetic categories, /w/ and flap.

4.1. Method

4.1.1. Subjects. Twenty-one native speakers of Korean (eight males and 13 females with a mean age of 23.4) participated in the experiment.

4.1.2. Stimuli. Speech materials consisted of nonsense words contrasting /r/ and /l/ in five positions. For each position there were two /r/-/l/ pairs, each with a different vowel, /a/ and /i/, as shown in Table 1. Two native speakers of American English (one male and one female) each produced these 16 words.

Table 1: Stimuli used in Experiment 3

Position	Stimuli
IP	/CV/ /ra/, /ri/, /la/, /li/
IC	/bCV/ /bra/, /bri/, /bla/, /bli/
IN	/VCV/ /ara/, /iri/, /ala/, /ili/
FC	/VCd/ /ard/, /ird/, /ald/, /ild/
FP	/VC/ /ar/, /ir/, /al/, /il/

4.1.3. Procedure. Subjects made cross-language goodness ratings of the stimuli. In each trial, Korean orthographies for Korean /w/ or flap were presented on a computer display, and one of the stimuli was presented auditorily over headphones. Subjects rated the "goodness-of-fit" of target consonant (/r/ or /l/) to the presented Korean category on the scale from 1 to 7 (1: least fit, 7: best fit). Each block of trials contained stimuli from one of the five syllable position conditions. Subjects completed five separate blocks, in a fixed ordering of syllable position conditions (IP, IC, IN, FC, and FP). Within each position block, eight stimuli (four syllables from two talkers) were presented twice in a random order. Subjects rated the goodness as Korean /w/ for one presentation of the stimulus, and as Korean flap for the other presentation. For randomly selected half of the stimuli, /w/ judgements were made first, and for the other half of the stimuli, flap judgements were made first.

4.2. Results and Discussion

Figure 4 shows averaged goodness ratings by position over subjects. For AE /l/'s in all positions and for AE /r/'s in four out of five positions (all except IN), goodness ratings as Korean flap were significantly higher than goodness ratings as Korean /w/. However, differences in goodness-of-fit rating scores between AE /r/ and /l/ differed significantly in IP, IC and IN positions, but not in FC or FP positions (Figure 5). In general, both /r/ and /l/ were assimilated into Korean flap, but /l/'s in IP, IC and IN positions better assimilated to Korean flap than /r/'s did, while /l/'s and /r/'s in FC and FP positions fitted about equally. Accordingly, in the context of PAM, we can categorize AE /r/-/l/ in IP, IC and IN positions as Category-Goodness Difference (CG) and /r/-/l/ in FC

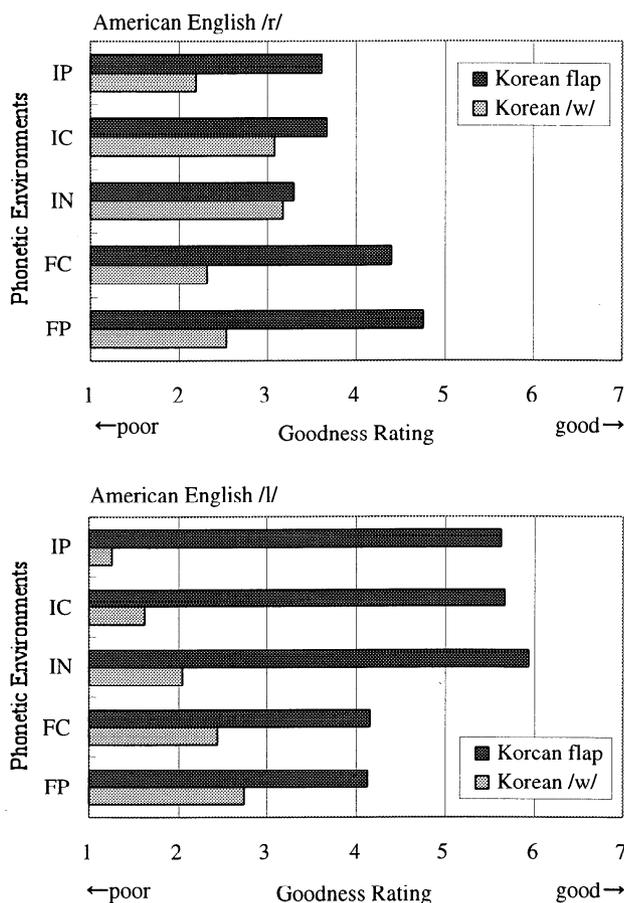


Figure 4: Averaged Goodness ratings by phonetic environments. The top panel indicates averaged goodness ratings of American English /r/ to Korean /w/ and flap, and the bottom panel indicates those of /l/.

and FP positions as Single-Category Assimilation (SC).

According to PAM, discrimination level is expected to be moderate to very good in a CG scenario, and poor in a SC scenario. From the above results, PAM predicts that Korean speakers have more difficulty distinguishing /r/ from /l/ in FC and FP than in IP, IC and IN. Note that Experiment 1 showed that Korean speakers' identification ability of /r/-/l/ in FC and FP positions was lower than in IP, IC and IN. We can conclude that the performance pattern of Korean speakers found in Experiment 1 is explainable by PAM.

5. GENERAL DISCUSSION

In Experiment 1, native speakers of Korean and Japanese showed different patterns of accuracy depending on the phonetic environment. For Japanese speakers, final cluster and final singleton are easier than other environments, which replicated the finding by Lively et al. [2]. In contrast, final cluster and singleton were shown to be most difficult for Korean speakers. Note that these two positions were not included in Ingram and Park's study [1]. We expanded their result in order to better understand language-specific position effect. Furthermore, Experiment 2 demonstrated that Korean speakers' responses were biased toward /l/. In Experiment 3

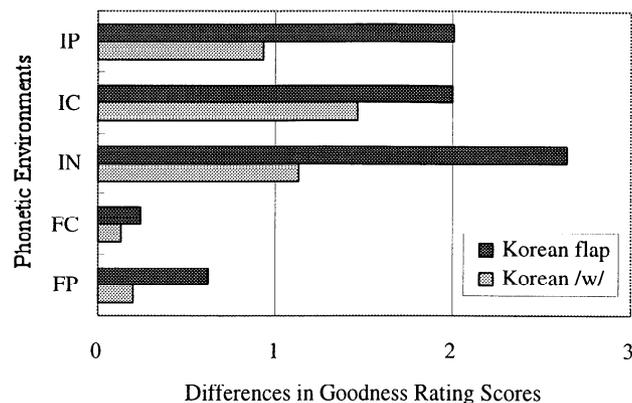


Figure 5: Differences in goodness-of-fit rating scores between AE /r/ and /l/

we assessed cross-language assimilation of AE /r/ and /l/ to Korean phonetic categories, and found that assimilation results predicted the Korean-specific perception pattern found in Experiment 1 fairly well. In order to further confirm this, experiments are currently underway to measure the assimilation of AE /r/ and /l/ to Japanese categories.

Since the interference of L1 seems to be a robust phenomenon in non-native speech perception, further research to understand the nature of this interference is necessary for making a practical contribution to L2 learning methodology.

ACKNOWLEDGMENTS

The authors wish to thank Prof. Youngsoo Kim at Ewha Womans University and Prof. Ilju Rha at Seoul National University for helping us run subjects in their laboratories. We also wish to thank Dr. Keiichi Tajima for his discussion and help in translation.

REFERENCES

- [1] Ingram, J.C.L and Park, S.-G., 1998. Language, context, and speaker effects in the identification and discrimination of English /r/ and /l/ by Japanese and Korean listeners. *The Journal of the Acoustical Society of America*, 103 (2), 1161-1174.
- [2] Lively, S.E., Pisoni, D.B., Akahane-Yamada, R., Tohkura, Y., and Yamada, T., 1994. Training Japanese listeners to identify English /r/ and /l/. III. Long-term retention of new phonetic categories. *The Journal of the Acoustical Society of America*, 96 (4), 2076-2087.
- [3] Best, C.T, 1995. A direct realist view of cross-language speech perception" In Strange, W. (ed.), *Speech Perception and Linguistic Experience*, 171-206, Baltimore: York Press.
- [4] 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 (4), 376-392.