DOES ALLOPHONIC KNOWLEDGE OF L1 CONTRIBUTE TO THE CORRECT DISCRIMINATION OF NON-NATIVE SOUNDS?

Shigeko Shinohara[a], *Qandeel Hussain*[b] & *Tomohiko Ooigawa*[a]

[a]Phonetics Laboratory, Sophia University, Tokyo, Japan;
[b]ARC Centre of Excellence in Cognition and its Disorders, Macquarie University, Sydney, Australia
shinoh-s@sophia.ac.jp, qandeel.hussain@students.mq.edu.au, ooigawaferchichi@gmail.com

ABSTRACT

The present paper investigated the perception of four Punjabi liquids /ɾ ɽ l ɭ/ and two stops /d ɖ/ by Japanese listeners. It has been reported that Japanese listeners have difficulty in perceiving the two-way liquid contrast of English because Japanese has only one liquid phoneme. Japanese /r/ has, in fact, a rich allophonic inventory. However, no detailed perception studies have been conducted to test whether rich allophonic inventory of Japanese /r/ helps Japanese listeners to perceive the phonemic four-way liquid and two-way stop contrasts of Punjabi. Our results from the forty-eight Japanese listeners suggested that, regardless of rich allophonic inventory of /r/, the liquid-liquid contrasts are still hard to discriminate for the Japanese listeners, but liquid-stop contrasts are easy to discriminate. On the other hand, retroflex stop [ɖ], an allophone of Japanese /r/, has given some advantage in the discrimination of /d-ɖ/ contrast and some disadvantage in liquid-stop contrasts.

Keywords: Japanese, Punjabi, perception, liquids, retroflex and dental.

1. INTRODUCTION

1.1. Background

Cross-linguistic studies on the non-native speech perception have shown that learning a new phonological contrast is always challenging for the L2 learners whose first language lacks that particular contrast. For instance, Japanese does not contrast /r/ and /l/. English, unlike Japanese, contrasts /r/ from /l/. Substantial research has been conducted on the perception of English /r/ and /l/ by Japanese listeners [4, 11]. The results of these studies suggest that Japanese listeners have difficulty in discriminating the English /r/ and /l/ contrast. This is further supported by neural imaging studies [20] showing that Japanese listeners are less sensitive to the English /ra/-/la/ contrast, compared to American English listeners. The aim of the present study was to look beyond the perception of traditional two-way liquid contrast of English by Japanese listeners. As Japanese has a rich allophonic inventory of liquids (retroflex and non-retroflex liquids and stops), this raises the question of whether allophonic knowledge of Japanese listeners help them to discriminate the speech sounds of a language (Punjabi) that has a rich phonemic inventory of retroflex and non-retroflex liquids and stops.

1.2. Phonetics and phonology of Japanese /r/

In Japanese phonology, /r/ displays an interesting pattern. It can occur before all of the five vowels and before a palatal glide /j/, the latter resulting in a palatalized /r/. The only consonant that can precede /r/ is the moraic nasal /N/, whose place assimilates to the following sound. /r/ is the fifth most frequent consonant in token frequencies among 15 consonant types across positions [9]. Its distribution is restricted to the word-medial position in native (Yamato) and mimetic lexicon, while it can freely occur word-initially in Sinitic words. Our preliminary findings of the segment frequencies in the Corpus of Spontaneous Japanese (CSJ) [2] indicate that Japanese /r/ is highly frequent compared to the dental stop /d/. In an intervocalic context, the sequence of moraic nasal and liquid (/Nr/) is remarkably less frequent than /r/ alone.

The phonetic studies of Japanese /r/ have reported a wide range of allophonic variations. [ɾ] is a basic realization of the Japanese /r/ [9, 10]. Other contextual, social and regional variants are [l], [ɭ], [ɮ], [r], [ɽ], [d] and [ɖ] [1, 12, 17]. However, those apparitions are hardly predictable, as vocalic contexts are not totally determinant. After a moraic nasal /N/, on the other hand, /r/ is mostly realized as a voiced apical stop [2]. Because of these allophonic variations, it is difficult to propose a single unified acoustic cue that distinguishes Japanese /r/ from other consonants.

1.3. Past studies

Last two or three decades have seen a plethora of literature related to the perception of English liquids by Japanese [11, 18], Korean [3] and Cantonese speakers [7]. Perception of liquid contrasts in other languages by Japanese listeners does not always follow the same patterns as those in English. Japanese listeners discriminate French two-way (/ʁ/ and /l/) liquid contrast better than English liquid contrast [5, 19]. As to Spanish three-way liquid (/r/, /ɾ/, and /l/) contrast, /r/-/ɾ/ and /r/-/l/ are also discriminated much better than English liquid contrast, but the listeners poorly perceive Spanish /ɾ/-/l/ contrast [14]. Thus, acoustic characteristics of the target L2 sounds play a significant role for categorisation, even though Japanese has only one liquid phoneme.

Another perception study of four-way liquid contrast of Russian liquids (palatalized and non-palatalized) by English, Japanese, Korean, Cantonese and Russian speakers shows that Japanese listeners perform similar to the native Russian listeners in perceiving the /l/ vs. /lʲ/ and /r/ vs. /rʲ/ contrasts but are less accurate in discriminating the /l/ vs. /r/ and /lʲ/ vs. /rʲ/ contrasts [16]. This is because Japanese has a contrastive set of plain and palatalized liquids.

As Japanese has a number of allophonic realizations of /r/, in the current study, we investigated whether allophonic knowledge of Japanese listeners can enhance their perceptual sensitivity towards the four-way liquid and two-way stop contrasts. To test this hypothesis, we selected Punjabi as a potential language that can contrast four liquids (dental tap /ɾ/ vs. retroflex flap /ɽ/, dental lateral /l/ vs. retroflex lateral /ɭ/) and two stops (dental /d/ vs. retroflex /ɖ/). Stops were added to better understand the perception of liquids. Based on the previous literature [11, 15] we predicted that Japanese listeners would have difficulty in perceiving the rhotic-lateral contrast but would perform better in the perception of /d-ɖ/ contrast since [d] and [ɖ] belong to distinct phonemes in Tokyo Japanese, /d/ and /r/, respectively. In other words, retroflex stop [ɖ] might be confused either with rhotics or with the dental stop by its phonetic characteristics. Furthermore, we predicted that Japanese listeners can easily discriminate the liquid-stop contrasts because Japanese contrasts liquid from stops.

2. METHOD

2.1. Participants

Forty-eight monolingual Japanese speakers residing in Tokyo area with no speaking or hearing impairment (18-23 years old) participated in the experiment. All the listeners were Japanese university students, had never stayed outside Japan for more than one month, and had never studied any foreign languages other than English.

2.2. Stimuli

2.2.1. Recording

The stimuli were consisted of six disyllabic pseudo-words, with the target consonants in word-medial position. As Japanese /r/ is more frequent in word-medial position [2], the same word-medial intervocalic context was selected. The stimuli are listed in Table 1. The effect of different vocalic environments on the retroflex consonants has been reported in the literature [13]. For the current study, we are only investigating the perception of liquids and stops between /a/ vowels.

Table 1: Stimuli used in the discrimination task.

	Dental	Retroflex
Rhotics	/ˈpʰaɾa/	/ˈpʰaɽa/
Laterals	/ˈpʰala/	/ˈpʰaɭa/
Stops	/ˈpʰada/	/ˈpʰaɖa/

The stimuli were recorded by a male Punjabi speaker (27 years old) in a sound attenuated room at the University of Agriculture, Pakistan. The utterances were recorded onto a Zoom digital voice recorder with an in-built microphone and digitized at 44.1 kHz with 16 bits. The peak amplitude of each stimulus was scaled to 0.99. For each word, two tokens were selected from the recorded materials as target stimuli for the current experiment.

2.2.2. Acoustic analyses of the stimuli

After the recording, we conducted acoustic analyses of the target stimuli. Dental coronals generally show raised F2 and flat F3 near the offset of the preceding vowel. On the other hand, retroflex coronals are characterised by the F2-F3 convergence near the offset of the preceding vowel and some lowering effect on F4 [6]. Table 2 presents the mean F2, F3 and F4 frequencies of the stimuli from mid and near the offset of the preceding /a/ vowel.

Table 2: Mean formant frequencies (Hz) of the Punjabi stimuli.

	F2	F3	F4
	Mid-Offset	Mid-Offset	Mid-Offset
/ɾ/	1089-1343	2518-2459	3428-3861
/ɽ/	1130-1314	2323-1575	3232-2944
/l/	1109-1322	2506-2566	3424-3846
/ɭ/	1125-1283	2251-1808	3099-2826
/d/	1126-1307	2529-2534	3636-3816
/ɖ/	1123-1543	2452-1674	3264-2784

As can be seen in Table 2, the F3 and F4 frequencies of the retroflex liquids /ɽ ɭ/ and stop /ɖ/ are significantly lowered near the offset, suggesting a clear retroflex articulation. F2 is raised and does not differ significantly among dentals and retroflex liquids, but in the case of retroflex stop /ɖ/, F2 is significantly raised near the offset.

2.3. Procedure

The perception experiment was conducted in a Computer-Assisted Language Learning classroom at a university in Japan. Praat was used as an interface. The stimuli were presented on a Windows PC with CZ530-A headphones.

The participants were asked to complete an AX discrimination task that included 168 trials ((15 different pairs + 6 same pairs) × 8 AX combinations), resulted in a total of 8064 trials (168 trials × 48 listeners). In each trial of the task, the participants listened to a pair of two stimuli, that were always different tokens, but the first stimulus was either the same word as the second one, or a different word (i.e. AA, AB, BB, or BA). For instance, for the /l-d/ pair, the stimuli were /ˈpʰada/-/ˈpʰala/ and /ˈpʰala/-/ˈpʰada/ of two tokens for each trial. The listeners had to judge whether the two stimuli were the same word or not, and had to click the *same* or *different* buttons on the screen, accompanied by Japanese translations. The pairs were presented in a random order. The interval between the two stimuli (i.e. between A and X) was 500 ms. The mean duration of each AX trial was 2000 ms. The reaction times of the each trial were also recorded. We asked the listeners to respond as fast as possible. In order to avoid the detection of the variations of /ˈpʰa/, we instructed the listeners to pay attention to the sound after /ˈpʰa/.

3. RESULTS

3.1. Discrimination

Figure 1 shows the results of the accurate discrimination of 15 different pairs. The pairs can be divided into three groups: *easy pairs* (/ɭ-d/, /l-d/, /ɾ-d/, /ɽ-d/, /ɭ-ɖ/, /ɾ-ɖ/, /ɽ-ɖ/ and /l-ɖ/) which show above 78% correct responses; *difficult pairs* (/ɭ-ɾ/, /l-ɽ/, /ɭ-ɽ/, /l-ɾ/, /l-ɭ/, and /ɾ-ɽ/) that are below 25%, and the *intermediate pair* (/d-ɖ/) that is around 50%.

According to a one-way repeated-measures ANOVA ($F\ (20, 940) = 164.941$, $p<.01$) and multiple comparisons (Bonferroni), the differences among the three groups were significant ($p<.01$). The differences among the pairs in the *difficult pairs* (i.e. liquid-liquid pairs) are not significant from each other ($p>.05$). In the *easy pairs* (liquid-stop pairs) the differences between the two top pairs in Figure 1 (/ɭ-d/ and /l-d/) and the two bottom pairs (/l-ɖ/ and /ɽ-ɖ/) were significant ($p<.01$). As predicted, the results indicate that the Japanese listeners poorly discriminated all the Punjabi liquid-liquid contrasts but performed very well in discriminating the Punjabi liquid-stop contrasts. Especially, they performed perfectly in the task of liquid versus dental stop contrasts (/ɽ-d/, /ɾ-d/, /l-d/ and /ɭ-d/), while they performed fairly well in discriminating the liquid versus retroflex stop contrasts (/l-ɖ/, /ɽ-ɖ/, /ɾ-ɖ/ and /ɭ-ɖ/). The result of /d-ɖ/ was different from both liquid-liquid contrasts and liquid-stop contrasts as being around 50%.

Figure 1: Correct rates of the different pairs.

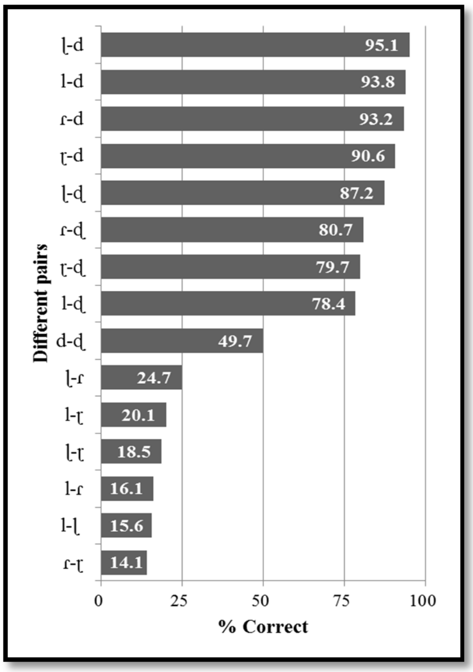
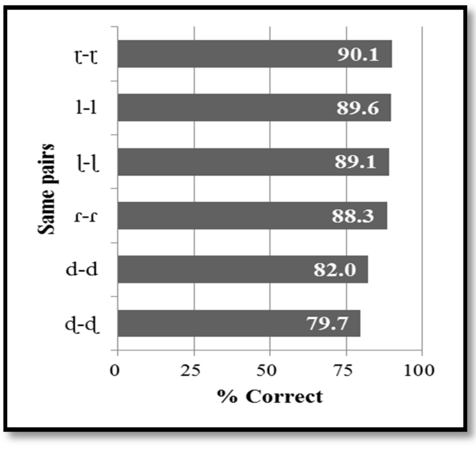


Figure 2 shows the correct response rates of the six same pairs. All the rates are relatively high and the ANOVA and the multiple comparisons showed no significant differences among the same pairs ($p>.05$).

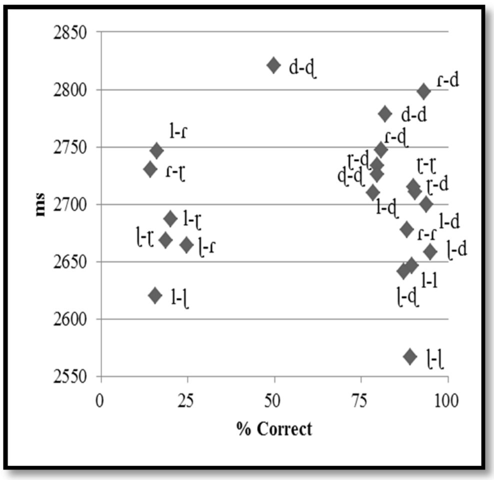
Figure 2: Correct rates of the same pairs.



3.2. Reaction time

Figure 3 shows the reaction time (x-axis) and the mean correct rates (y-axis) of the discrimination task. Each reaction time was measured from the beginning of the first stimulus of each trial. Out of total 8064 trials, we excluded 243 trials as outliers (too fast: 131, and too late: 112). The definition of *too fast* responses was that the time that is shorter than 2000 ms because the mean value of the total length of each AX trial was about 2000 ms long. The definition of *too slow* responses was that the time that is more than 3 standard deviations longer from the mean [8].

Figure 3: Correct rates (x-axis) and reaction time (y-axis) of the discrimination task.



There is no correlation between the correct rates and the reaction times of the discrimination task (Pearson's r = .018). However, the reaction time of the /d-ɖ/ contrast was longer compared to other contrasts. According to a one-way repeated-measures ANOVA (F (20, 940) = 6.835, p<.01) and multiple comparisons (Bonferroni), the differences between the reaction times of the top /d-ɖ/ and the bottom (/ɭ-ɭ/, /l-ɭ/, ɭ-ɖ/, /l-l/, /ɭ-d/, /ɭ-ɾ/) contrasts are significant (p<.01). The listeners might have hesitated to discriminate the /d-ɖ/ contrast compared to some other pairs which resulted in a longer reaction time.

4. DISCUSSION

The present paper investigated the perception of four-way liquid and two-way stop contrasts of Punjabi by Japanese listeners. The overall results indicated that liquid-stop contrasts were easy to discriminate, compared to the liquid-liquid contrasts. There were significant differences among the correct discrimination rates of the Punjabi liquid-liquid contrasts (less than 25%) and those of the liquid-stop contrasts (more than 78%). These results suggested that the listeners assimilated the four Punjabi liquids to the native single category, Japanese /r/. On the other hand, the listeners were able to discriminate the differences between the liquids and stops. As to the stop-stop contrast, /d-ɖ/, the listeners perceived the difference at the rate of about 50%, which was significantly different from those for the liquid-liquid and the liquid-stop contrasts. This pattern of perception emerges because Japanese contrasts liquids and stops but does not have a contrastive liquid inventory.

Studies on the perception of Hindi /d-ɖ/ contrast [e.g. 15] reported that Japanese listeners perform better than English speakers in /d-ɖ/ discrimination because Japanese contrasts dental stop /d/ and tap /r/. As noted earlier, retroflex [ɖ] is one of the allophonic realizations of Japanese /r/. The listeners might have assimilated the Punjabi dental and retroflex stops to Japanese /d/, and sometimes they assimilated Punjabi /ɖ/ to Japanese /r/.

Looking more closely at the *easy* liquid-stop pairs, those involving dental stop (/ɽ-d/, /ɾ-d/, /l-d/ and /ɭ-d/) tend to show higher correct rates compared to the liquid-stop pairs involving retroflex stop (/l-ɖ/, /ɽ-ɖ/ /ɾ-ɖ/ and /ɭ-ɖ/). If listeners identified [ɖ] as Japanese /r/, that would explain the better discrimination performance in the dental stop versus liquid contrasts than the retroflex stop versus liquid contrasts. Since both liquids and [ɖ] can be identified as single /r/, it would lead to less accurate discrimination scores. These gradual performance differences might be due to individual discrimination patterns or to the more limited occurrence of [ɖ] as a contextual allophone of /r/ in L1, i.e. frequency of allophones might influence the identification patterns.

Another point worth pursuing would be the role of phonetic similarity among the flap/tap and stops. We also noted that the listeners took significantly longer time to discriminate one of the rhotic-stop pairs, /ɾ-d/, compared to two of the lateral-stop pairs, /ɭ-d/ and /ɭ-ɖ/ (p<.01), among the similar pairs showing weaker but similar tendencies. This pattern suggests that while tap/flap and voiced stop coronals are confusable to some extent, Japanese speakers be trained somehow in their L1 to discriminate tap/flap and voiced coronal stop.

We concluded that regardless of rich allophonic inventory of /r/, for the Japanese naïve listeners, the discrimination of all the Punjabi liquid contrasts is difficult, that of liquid-stop contrasts is easy but the difference is gradual and reflects the role of L1 allophonic variation.

5. REFERENCES

[1] Arai, T. 2013. On why Japanese /r/ sounds are difficult for children to acquire. *Proc. 14th INTERSPEECH* Lyon, 2445-2449.

[2] Furui, S. 2008. *The corpus of spontaneous Japanese, second printing, Konotoha-1 (DVD database)*. Tokyo: NINJLA and NIICT.

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

[4] Goto, H. 1971. Auditory perception by normal Japanese adults of the sounds "L" and "R". *Neuropsychologia* 9 (3), 317-323.

[5] Hallé, P. 2000. Traitement phonétique et reconnaissance des mots chez le jeune enfant et chez l'adulte. HDR thesis, Paris 5 René Descartes, Paris.

[6] Hamann, S. 2003. *The phonetics and phonology of retroflexes*. Utrecht: LOT.

[7] Henly, E., Sheldon, A. 1986. Duration and context effects on the perception of English /r/ and /l/: A comparison of Cantonese and Japanese speakers. *Language Learning* 36 (4), 505-522.

[8] Jiang, N. 2012. *Conducting reaction time research in Second language studies*. New York: Routledge.

[9] Labrune, L. 2012. *The phonology of Japanese*. Oxford: Oxford University Press.

[10] Labrune, L. 2014. The phonology of Japanese /r/: a panchronic account. *Journal of East Asian Linguistics* 23 (1), 1-25.

[11] MacKain, K. S., Best, C. T., Strange, W. 1981. Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics* 2 (04), 369-390.

[12] Magnuson, T. J. 2008. *What /r/ sounds like in Kansai Japanese: a phonetic liquid variation in unscripted discourse.* MA thesis, University of Victoria, British Columbia.

[13] Ohala, M., Ohala, J. J. 2001. Acoustic VC transitions correlate with degree of perceptual confusion of place contrast in Hindi. In: Grønnum, N., J. Rischel (eds), *Travaux du Cercle Linguisticque de Copenhague*. Copenhagen: Reitzel, 265-284.

[14] Ooigawa, T. 2013. Perception of /l/ and /r/ in English, French and Spanish by Japanese listeners. *Sophia Linguistica* 61, 49-68.

[15] Pruitt, J. S., Jenkins, J. J., Strange, W. 2006. Training the perception of Hindi dental and retroflex stops by native speakers of American English and Japanese. *The Journal of the Acoustical Society of America* 119 (3), 1684-1696.

[16] Smith, J., Kochetov, A. 2009. Categorization of non-native liquid contrasts by Cantonese, Japanese, Korean, and Mandarin listeners. *Toronto Working Papers in Linguistics* 34, 1-15.

[17] Sugai, K. 2002. *Realization of /r/ sound in Standard Japanese: An acoustic and phonological analysis.* MA thesis, Sophia Univ., Tokyo.

[18] Yamada, R. A., Tohkura, Y. 1992. The effects of experimental variables on the perception of American English /r/ and /l/ by Japanese listeners. *Percept Psychophys* 52 (4), 376-392.

[19] Yamasaki, H., Hallé, P. A. 1999. How do native speakers of Japanese discriminate and categorize French /r/ and /l/? *Proc. 14th ICPhS* San Francisco, 909-912.

[20] Zhang, Y., Kuhl, P. K., Imada, T., Kotani, M., Tohkura, Y. 2005. Effects of language experience: neural commitment to language-specific auditory patterns. *Neuroimage* 26 (3), 703-720.