# Japanese Listeners' Perception of American English Intervocalic /l/, /r/ and Flapped /t/

Tomohiko Ooigawa

Sophia University, Japan,
*ooigawaferchichi@gmail.com*

## Abstract

In order to apply the principles of perception to EFL phonetic education, the present study examined perception of American English intervocalic /l/ ([l]), /r/ ([ɹ]) and flapped/tapped /t/ ([ɾ]) by Japanese listeners. The results indicated that the listeners discriminated the [ɹ]-[ɾ] contrast better than the [l]-[ɹ] and [l]-[ɾ] contrasts. Most of both English [l] and [ɹ] stimuli were perceptually assimilated to Japanese /r/ or sounds that include /r/, while [ɾ] was mostly assimilated to Japanese /d/ and /r/. English [ɾ] was rated as the most Japanese-like sound of the three consonants.

**Keywords:** Perception, flap/tap, liquids, Japanese.

## 1. Introduction

Intervocalic /t/ and /d/ can be realized as [ɾ] (a flap or tap) in North American English (e.g., [1, 2, 3, 4]), which is traditionally referred to as "flapping." The typical examples are the realization of /t/ and /d/ in *Sa**t**ur**d**ay*. In Japanese, the sole liquid phoneme can be also realized as [ɾ], though it has various allophones, including [l] (e.g., [5, 6, 7]). For example, /r/ in /ka**r**ate/ (空手) can be realized as [ɾ]. Many studies on the perception of English /l/ and /r/ have been carried out and indicated that, generally, Japanese listeners poorly discriminated the contrast (e.g., [8, 9, 10]), because the listeners perceptually assimilated English /l/ and /r/ sounds to Japanese /r/ in general (e.g., [11, 12, 13]). However, to my knowledge, there are only three studies that deal with the perception of English [ɾ] by Japanese listeners. The first study ([14]) examined the English-to-Japanese *hiragana* syllabary transcription, and indicated that Japanese listeners perceptually assimilated English [ɾ] to Japanese /r/ or /d/ or /t/ (mostly to /r/). The second study ([15]) pointed out that English [ɾ] often made it difficult for Japanese EFL learners to recognize words that include [ɾ]. For instance, it is difficult to recognize "letters" "waiting" and "item," "even if they know these words and use them in *katanaka*" ([15]). The third study ([16]) examined the English-to-English alphabet identification, and suggested that Japanese listeners confused English [l] and [ɾ] more often than [ɹ] and [ɾ]. The suggested reason for the confusion was that both [l] and [ɾ] are allophones of Japanese /r/ ([16]). The suggestion implied that the two sounds are close to each other for Japanese listeners, while [ɹ] is less common than [l] and [ɾ] as an allophone of Japanese /r/, so that [ɹ] can be perceived as a different sound from [l] and [ɾ] to some extent.

An identification test of English [ɾ] was carried out in [14], yet the stimuli did not include the liquids. The second study ([15]) carried out a dictation test, and the third one ([16]) did not examine perceptual assimilation of English [ɾ] by Japanese listeners. Thus, so far, no studies have examined Japanese listeners' discrimination by using a discrimination test, goodness ratings and comparisons of perceptual assimilation of English [l], [ɹ] and [ɾ]. Therefore, in order to identify the perception in more detail, the present research conducted discrimination and goodness rating tests, carrying out another English-to-Japanese *katakana* syllabary identification test of American English [l], [ɹ] and [ɾ]. The final goal of the present research is to apply the findings to phonetic education of Japanese EFL learners.

## 2. Experiments

### 2.1. Procedures

The present study conducted three kinds of tests: discrimination, identification and goodness rating tests, and the target stimuli and the listeners were identical. None of the speakers and listeners had any perceptual and physical difficulties participating in the recording or the tests. The experiments were conducted in a Computer-Assisted Language Learning (CALL) classroom in a university in Japan. Praat (version 5.3.77) ([17]) was used as an interface. Prior to the tests, the volume was adjusted to a comfortable listening level. The listeners completed each task while wearing headphones (CZ530-A) in front of the computer. The following procedure was common to the three tasks; The stimuli were


10.21437/ISAPh.2016-13

presented in random order to each listener over headphones from the computer. The listeners were allowed to make a correction to their response before clicking the *Next* button to proceed to the next trial. Once they clicked the *Next* button, they were not allowed to return to the previous trial.

### 2.2. Stimuli and listeners

Two native speakers of American English who came from the United States individually participated in the recording. One was a 40-year-old man from Hartford, CT, and the other was a 21-year-old woman from Buffalo, NY (hereafter, Speaker 1 and Speaker 2, respectively). The speakers had grown up in a monolingual family. The results obtained from the two speakers' stimuli were combined in order to avoid complexity. The target stimuli were different for each speaker (see, Table 1). In common, the stimulus materials consisted of bi-syllabic trochaic English words with intervocalic /l/ or /r/ or /t/, which had a bilabial plosive followed by /ɛ/ at the onset in the first syllable, and had /i/ in the second unstressed syllable. The carrier sentence for Speaker 1 was *Repeat* "__" *twice*, and that for Speaker 2 was *Say* "__" *please*. The speakers read out the target words with the carrier sentences along with some distracter words in random order at least ten times in a soundproof room. The utterances were recorded onto a digital recorder (PCM-M10) through a microphone (ECM-MS957) and digitized at 48 kHz with 16 bits. The target word tokens were extracted from the carrier sentences. For each target word, three tokens were selected from the recorded materials as stimuli.

Figures 1 and 2 show the waveforms and spectrograms of one token of the target stimuli displayed by the software program Praat [17]. In order to show the acoustic data, each token of the stimuli was extracted from the first zero-crossing of the burst of the first plosive (i.e., the first spike in the waveform). The durations were aligned to 300 ms. The acoustic data of the stimuli in the present study is consistent with those in the acoustic studies ([2, 18, 19, 20]); The most important acoustic difference between English /l/ and /r/ is in the frequency of the third formant. The third formant of English [l] is higher than that of English [ɹ]. The flap [ɾ] has a very short closure.

The listeners were 34 students (18-22 years old, mean: 18.9) of a university in Japan, and they asserted that their native language was Japanese and that they were not bilinguals in the questionnaire. The listeners had never stayed outside Japan for more than one month, had never majored in linguistics or a foreign language and had never studied linguistics or foreign languages other than English.

**Table 1:** Word list of the target stimuli.[1]

| Speaker 1: | | | Speaker 2: | | |
|---|---|---|---|---|---|
| belly | /ˈbɛli/ | [ˈbɛli] | Pelly | /ˈpɛli/ | [ˈpʰɛli] |
| berry | /ˈbɛri/ | [ˈbɛɹi] | Perry | /ˈpɛri/ | [ˈpʰɛɹi] |
| Betty | /ˈbɛti/ | [ˈbɛɾi] | petty | /ˈpɛti/ | [ˈpʰɛɾi] |

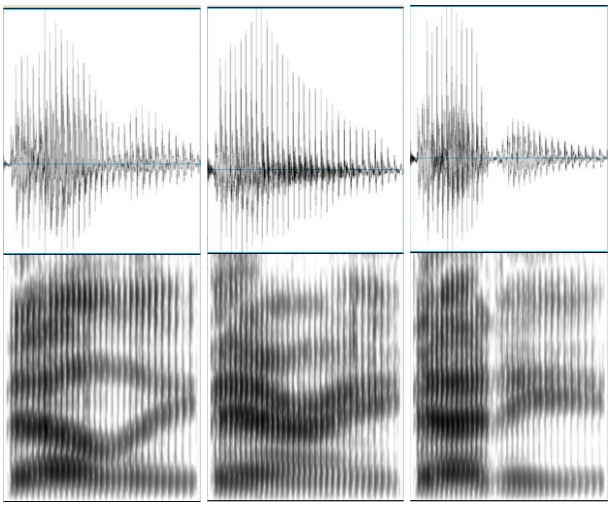

**Figure 1:** The waveforms (upper) and spectrograms (lower) of a token of [ˈbɛli], [ˈbɛɹi] and [ˈbɛɾi] produced by Speaker 1 (duration: 300 ms, view range: 0.0-5000.0 Hz, window length: 0.005 s, dynamic range: 50.0 dB.).

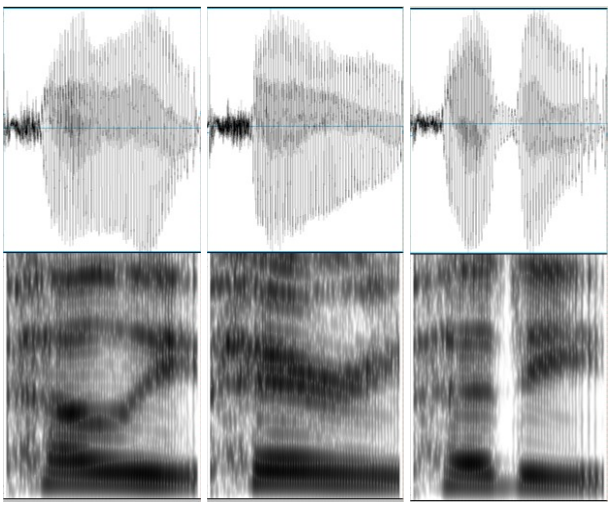

**Figure 2:** The waveforms (upper) and spectrograms (lower) of a token of [ˈpʰɛli], [ˈpʰɛɹi] and [ˈpʰɛɾi] produced by Speaker 2. The settings are identical to those in the caption of Figure 1.

### 2.3. Discrimination test

The listeners completed an AXB discrimination task. In each trial, the participant listened to three stimuli (i.e., AAB, ABB, BBA or BAA) that were different tokens, but the second stimulus was the identical word to the first, or to the third stimulus (e.g., *belly-belly-berry* or *belly-berry-berry*). The listeners judged whether the second stimulus was more

similar to the first, or to the third stimulus, and they clicked the *First* or *Third* button on the screen. The test included 80 trials: (12 AXB triplets × 3 contrasts × 2 speakers) + 8 distracters. The distracters were contrasts of *night vs. zoo*, *thank you vs. zoo*, *cafeteria vs. kangaroo* and *media vs. vanilla* (2 trials each). The inter-stimulus intervals were 100 ms.

*2.3.1. Results of the discrimination test*

Figure 3 shows the mean correct rates obtained by the discrimination test. The listeners discriminated the [ɹ]-[ɾ] contrast (92%) better than the contrasts of [l]-[ɹ] (84%) and [l]-[ɾ] (85%). A one-way repeated-measures ANOVA ($F(2, 66) = 15.971$, $p < .001$) and the multiple comparisons (Bonferroni) showed that there were significant differences in the scores between [l]-[ɹ] and [ɹ]-[ɾ] ($p < .01$), and between [l]-[ɾ] and [ɹ]-[ɾ] ($p < .01$), and that there was no significant difference in the scores between [l]-[ɹ] and [l]-[ɾ] ($p = 1.000$).

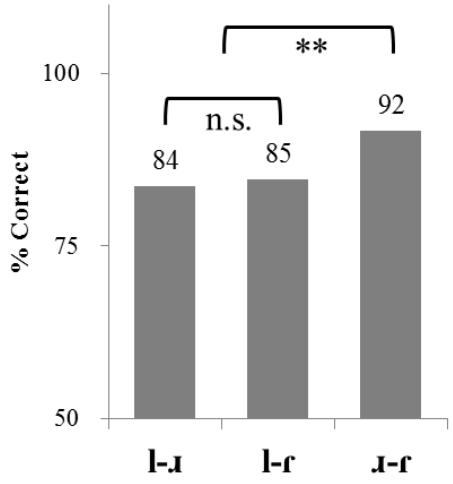


**Figure 3:** Mean correct discrimination rates (%).

## 2.4. Identification test

After the discrimination task, the listeners completed a forced-choice identification task with the goodness rating task (see, 2.5). In each trial, the participant listened to a stimulus and judged which button on the screen indicated the most similar Japanese transcription for the stimulus. When one play was not sufficient to make a choice, they were permitted to replay the recording up to three times by clicking the *One more* button for each trial. The identification test of the present study included 36 trials: 3 words × 3 tokens × 2 speakers × 2 repetitions. On the screen of the computer, seven choices written in the Japanese *katakana* syllabaries were presented: "ベリ /beri/," "ベウリ /beuri/," "ベルリ /beruri/," "ベッリ /beQri/," "ベウィ /bewi/," "ベディ /bedi/," "ベティ /beti/" for the trials of the stimuli produced by Speaker 1, and "ペリ /peri/," "ペウリ /peuri/," "ペルリ /peruri/," "ペッリ /peQri/," "ペウィ /pewi/," "ペディ /pedi/" and "ペティ /peti/" for the trials of the stimuli produced by Speaker 2. The transcriptions were selected from ones obtained in the preliminary free transcription test that was conducted under identical conditions to the present identification test except for the followings; Twelve participants listened to the stimuli and transcribed each stimulus in Japanese *katakana* syllabaries on a sheet of paper. The listeners were permitted to replay the recording up to five times for each trial.

*2.4.1. Results of the identification test*

Figure 4 shows each perceptual assimilation proportion obtained by the identification test. The assimilation patterns of [l] and [ɹ] are much closer than the pattern of [ɾ]. The listeners chose "リ /ri/ (53%)," "ウリ /uri/ (17%)," "ルリ /ruri/ (11%)," "ディ /di/ (5%)," etc. for [l], while they chose "リ /ri/ (43%)," "ウリ /uri/ (26%)," "ルリ /ruri/ (14%)," "ディ /di/ (1%)," etc. for [ɹ]. They chose "ディ /di/ (70%)," "リ /ri/ (16%)," "ティ /ti/ (9%)," etc. for [ɾ].

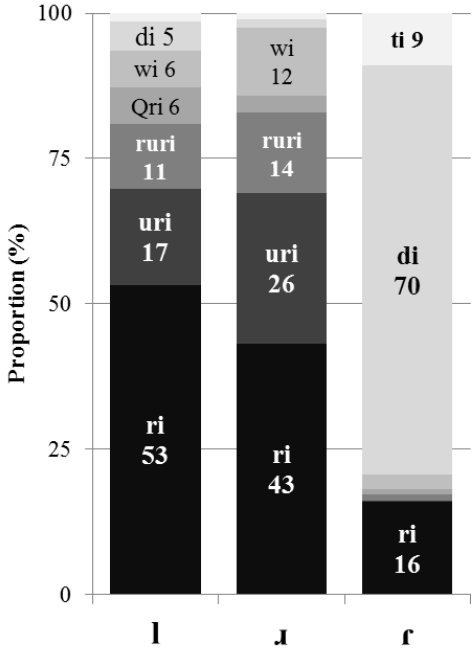


**Figure 4:** Identification proportion (%).

## 2.5. Goodness rating test

The goodness rating test was conducted in parallel with the identification test (see, 2.4). In addition to the Japanese transcriptions, the five-point scale goodness rating choices were presented on the screen. In each trial, the participants listened to a stimulus and judged how close the stimulus was to the Japanese sound they chose. The score *5* indicated the closest sound to the Japanese sound, while the

score *1* indicated the least close sound to the Japanese sound.

*2.5.1. Results of the goodness rating test*

Figure 5 shows the mean scores obtained by the goodness rating test. The listeners rated [ɾ] (3.3) higher than [ɹ] (2.8) and [l] (3.0). A one-way repeated-measures ANOVA ($F(2, 66) = 8.380$, $p < .01$) and the multiple comparisons (Bonferroni) showed that there were significant differences in the scores between [ɹ] and [ɾ] ($p < .01$), and between [l] and [ɾ] ($p < .05$), and that there was no significant difference in the scores between [l] and [ɹ] ($p = .424$).

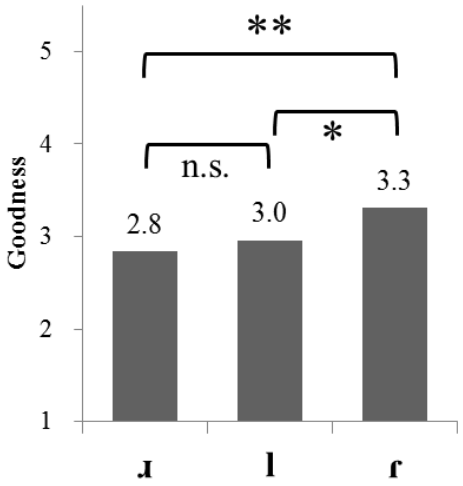


**Figure 5:** Mean scores of the goodness rating test. The score *1* indicates the least Japanese-like and *5* indicates the most Japanese-like.

## 3. Discussion

### 3.1. Summary of the results

Figure 6 summarizes the results in the present study. The results of the discrimination test indicated that the monolingual Japanese listeners discriminated the American English [ɹ]-[ɾ] contrast significantly better than the contrasts of [l]-[ɹ] and [l]-[ɾ]. The discrimination of the contrast of [l]-[ɾ] was as difficult as that of [l]-[ɹ] for the listeners. It was consistent with the results in the previous study ([16]) that the listeners discriminated [l]-[ɾ] worse than [ɹ]-[ɾ] in the present study, because the study ([16]) indicated that Japanese listeners confused English [l] and [ɾ] more often than [ɹ] and [ɾ].

The results of the identification test indicated that the monolingual Japanese listeners generally assimilated the American English [l] and [ɹ] stimuli to Japanese /r/, and the [ɾ] stimuli to Japanese /d/. The perceptual assimilation patterns of English liquids shown in the present study were consistent with those in the previous studies ([11, 12, 13]). As to the assimilation patterns of [ɾ], the result in the present study was consistent with that in the previous study ([14]). However, we can find a difference between the two studies in the assimilation proportion of Japanese /r/ and /d/. English [ɾ] was assimilated more to Japanese /r/ than to /d/ in [14], while the present study showed more assimilation to Japanese /d/ than to /r/. Yet, the two studies showed the perceptual assimilation patterns of English [ɾ] to Japanese /r/, /d/ and /t/ in common. It is possible that [ɾ] in the previous study was "more flapped" than the sound in the present study. We should use more naturally produced flap stimuli in order to replicate the studies in the future.

The results of the goodness rating test indicated that the monolingual Japanese listeners judged English [ɾ] was the closest sound to Japanese sounds of the three English consonants (i.e., [l], [ɹ] and [ɾ]). It was consistent with the suggestions in the previous study ([16]) that English [ɾ] was closer to a Japanese sound than English [ɹ]. Yet, it was inconsistent with the suggestions that both [l] and [ɹ] were rated equally. The study ([16]) implied that English [l] and [ɾ] are closer sounds to Japanese sounds than English [ɹ], because [l] and [ɾ] are more common allophones of Japanese /r/ than [ɹ].

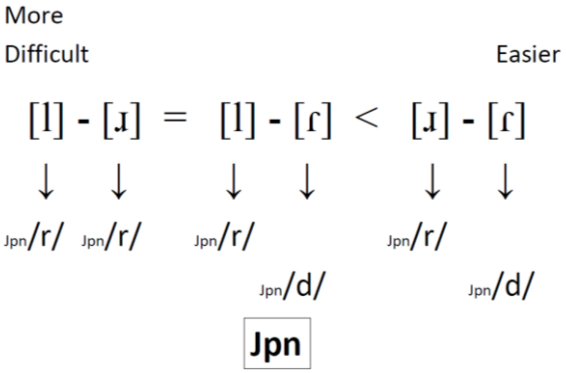


**Figure 6:** Schematic figure indicating the summary of the results.

### 3.2. Remaining questions

The discrimination of the English [l]-[ɹ] contrast was more difficult than that of [ɹ]-[ɾ], because both [l] and [ɹ] were assimilated to the same Japanese sound category (i.e., /r/), but [ɹ] and [ɾ] sounds were assimilated to different Japanese sound categories from each other (i.e., /r/ and /d/). However, the results of the identification test did not give us any answers to the question of why the discrimination of the English [l]-[ɾ] contrast was as difficult as that of the contrast of [l]-[ɹ], because the perceptual assimilation patterns are different in [l]-[ɾ], but identical in [l]-[ɹ]. Moreover, the results of the goodness rating test in the present study did not

answer the question either. As the goodness rating scores of [l] and [ɹ] are equal, the conditions of the [l]-[ɾ] and [ɹ]-[ɾ] are identical: "Japanese /r/-like, not close" vs. "Japanese /d/-like, close" (see, Figure 6).

The articulatory differences in the three consonants might account for the phenomena (see, Table 2). [l], [ɹ] and [ɾ] are voiced sounds, and the places of articulation are almost identical (i.e., coronal). [l] and [ɹ] share an identical manner of articulation "approximant," though [l] is lateral and [ɹ] is central. The tongue tip touches the alveolar ridge when producing both [l] and [ɾ]. The [l]-[ɹ] and [l]-[ɾ] contrasts share the three points, while [ɹ]-[ɾ] has only two commonalities, which means [ɹ]-[ɾ] is a more different sound pair than [l]-[ɹ] and [l]-[ɾ]. Therefore, the discrimination of [ɹ]-[ɾ] can be easier than [l]-[ɹ] and [l]-[ɾ].

**Table 2:** Ariticulatory differences of [l], [ɹ] and [ɾ].

| | [l] | [ɹ] | [ɾ] |
|---|---|---|---|
| Voicing | + | + | + |
| Place | Cor. | Cor. | Cor. |
| Manner | Ap. | Ap. | Flap |
| Contact | + | - | + |

As suggested in [16], the allophonic distribution of Japanese /r/ also might account for the phenomena. As allophones of the phoneme, [l] and [ɾ] are more common than [ɹ], so that [ɹ] can be perceived as a different sound from [l] and [ɾ] to some extent, resulting in better discrimination of [ɹ]-[ɾ] than those of [l]-[ɹ] and [l]-[ɾ]. Moreover, there is no Yamato (i.e., native) Japanese and Sino-Japanese words that include [di], which might be because of confusability with /ri/ (can be realized as [ɾi] and [li]). If so, it is reasonable that [li]-[ɾi] discrimination is more difficult than [ɹi]-[ɾi]. We should use stimuli consisting of other vowel contexts than /-i/ in order to confirm the suggestions in the future. However, another remaining question is why English [l]-[ɹ] discrimination is difficult for native speakers of Japanese. [ɹ] is a very unnatural realization as Japanese /r/, but it sounds close to Japanese /r/.

## 4. Conclusions

The present study examined Japanese listeners' discrimination, identification and goodness ratings of intervocalic /l/ ([l]), /r/ ([ɹ]) and flapped/tapped /t/ ([ɾ]) in American English. According to the results, the discrimination of the [ɹ]-[ɾ] contrast was easier than those of the [l]-[ɹ] and [l]-[ɾ] contrasts, which indicated that [l]-[ɹ] and [l]-[ɾ] were more perceptually confusable than [ɹ]-[ɾ] for native speakers of Japanese. Both [l] and [ɹ] sounds were perceptually assimilated to Japanese /r/ (i.e., simply, "リ /ri/") or sounds that include /r/ (e.g., "ウリ /uri/"), while [ɾ] was assimilated to Japanese /d/ and /r/ in most of the cases in the identification test. The results of the goodness rating test showed that [ɾ] was the most Japanese-like sound of the three consonants for native speakers of Japanese.

## 5. Educational implications

According to [21], the consonants /l/, /r/ and /t/ appear frequently in English. We can say that listeners often encounter the three sounds [l], [ɹ] and [ɾ] (e.g., in *Saturday, party people, check it out, Let it go/be, a lot of, rater* and *later*). In addition, it is often difficult for Japanese EFL learners to recognize words that include [ɾ] ([15]). The previous study ([16]) and the present study indicated that three consonants were perceptually confusable (esp., [l]-[ɹ] and [l]-[ɾ]). Therefore, English teachers should teach how to pronounce English [l], [ɹ] and [ɾ], and listening skills for the three consonants to Japanese EFL learners. However, as to [ɾ], we should give priority to listening. There are various kinds of Englishes: "Inner Circle" "Outer circle" and "Expanding Circle" Englishes ([22]). Some English accents have flapping and others do not ([1, 23]), even within Inner Circle Englishes. Therefore, I do not think that all Japanese EFL learners need to acquire flapping, because as a speaker, we may choose to produce one of the English accents (or Englishes).[2] However, we cannot choose pronunciation as a listener. We need to recognize [ɾ] as a realization of /t/, because we may encounter speakers whose English has flapping. We have to listen to utterances of various kinds of Englishes and understand them in the real world.

However, it is difficult to clearly separate teaching how to pronounce them and how to listen to them. Here I propose some teaching methods, including pronunciation instructions and listening skills based on the findings of the present study, which I practice in my English phonetics classes; Firstly, before teaching [ɾ], we have to confirm that learners have sufficiently acquired the English /l/-/r/ contrast. If they have learned it insufficiently, it is possible that they may confuse the three consonants and say, for example, *New York Silly* instead of *New York City*. As to the production of [ɾ], we should start with making learners pronounce [d] as /t/, and gradually move on to making them pronounce [ɾ] (e.g., [sɪti] → [sɪdi] → [sɪɾi] as *city*), because Japanese learners can produce [d] more easily than they pronounce [ɾ] intentionally,[3] and because [ɾ] is close to [d]. In addition, it is very effective to pronounce [lː], [ɹː] and [ɾ] (e.g., [bɛlːi], [bɛɹːi] and

[bɛɾi] as *belly, berry* and *Betty,* respectively) in order to emphasize the durational, articulatory and auditory differences among the three consonants. We can pronounce [l] and [ɹ] while lengthening the duration of the two consonants. As shown in 2.2, [ɾ] has a very short closure, and we cannot lengthen its duration; Long [ɾ] can be virtually identical to [d].

As to the perception, we have to teach that "t" is a very changeable sound; It can be [$t^h$] (strong "t"), [ʔ] (a sudden stop or virtually no sound), [d] and [ɾ] (e.g., in ***t**en, s**t**op, bu**tt**on, hi**t*** and *ma**tt**er*). And, we should recommend that students should not always expect clear t-like sounds for "t." In addition, because [ɹ]-[ɾ] is easier to discriminate than [l]-[ɾ], we should conduct discrimination trainings for [ɹ]-[ɾ] before those for [l]-[ɾ], so that the learners can more easily notice the auditory differences between [ɾ] and the liquids.

## 6. Acknowledgements

I would like to express my gratitude to the listeners and the native speakers who were willing to participate in the research. My thanks also go to Dr. Kinuko Takahashi of Sophia University and Prof. Ian Wilson of the University of Aizu. They gave me some precious insights for this study. Thank you very much.

## 7. References

[1] Wells, J. C. 1982. *Accents of English*. Cambridge: Cambridge University Press.

[2] Zue, V. W., Laferriere, M. 1979. Acoustic study of medial /t,d/ in American English. *J. Acoust. Soc. Am.* 66(4), 1039-1050.

[3] Derrick, D., Gick, B. 2011. Individual variation in English flaps and taps: A case of categorical phonetics. *The Canadian J. of Linguistics* 56(3), 307-319.

[4] Herd, W., Jongman, A., Sereno, J. 2010. An acoustic and perceptual analysis of /t/ and /d/ flaps in American English. *J. of Phonetics* 38(4), 504-516.

[5] Vance, T. J. 2008. *The Sounds of Japanese*. New York: Cambridge University Press.

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

[7] Sugai, K. 2002. Realization of /r/ Sound in Standard Japanese: An Acoustic and Phonological Analysis. MA thesis, Sophia Univ., Tokyo.

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

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

[10] Miyawaki, K., Jenkins, J. J., Strange, W., Liberman, A. M., Verbrugge, R., Fujimura, O. 1975. An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics* 18(5), 331-340.

[11] Komaki, R., Akahane-Yamada, R., Choi, Y. 1999. Effects of native language on the perception of American English /r/ and /l/: Cross-language comparison between Korean and Japanese [written in Japanese]. *Technical Report of IEICE* SP99-45, 39-46.

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

[13] Guion, S. G., Flege, J. E., Akahane-Yamada, R., Pruitt, J. C. 2000. An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *J. Acoust. Soc. Am.* 107(5), 2711-2724.

[14] Price, P. J. 1981. A Cross-Linguistic Study of Flaps in Japanese and in American English. Ph.D. thesis, University of Pennsylvania.

[15] Hayashi, K. 2009. Recognition of Flapped /t/ by Japanese University Students. *Chiba University of Commerce Review* 46(4), 15-34.

[16] Kitahara, M., Tajima, K., Yoneyama, K. 2015. Perception of flap/stop, and liquid sounds in American English by Japanese speakers (written in Japanese). *Proc. of the 29th General Meetings of the PSJ* Kobe Univ., 74-79.

[17] Boersma, P., Weenink, D. 2014. Praat: doing phonetics by computer [Computer program] (Version 5.3.77).

[18] Ladefoged, P. 2003. *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*. Malden, MA: Wiley-Blackwell.

[19] Ladefoged, P. 2006. *A Course in Phonetics* (5th ed.). Boston, USA: Thomson Wardsworth.

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

[21] Ladefoged, P. 2005. *Vowels and Consonants: An Introduction to the Sounds of Languages* (2nd ed.). Malden, MA: Blackwell.

[22] Kachru, B. B. 1992. *The Other Tongue: English across Cultures* (2nd ed.). Urbana/Chicago: University of Illinois Press.

[23] Wells, J. C. (Ed.) (2008) Longman Pronunciation Dictionary (3rd ed.). Edinburgh, UK: Pearson Education.

[1] In order to avoid confusion, the distinctions between [l] and [ɬ], between [ɹ] and [ɻ], and between tap and flap were ignored in this paper.

[2] It is controversial which English variant(s) we should choose as (a) model(s) for EFL learners.

[3] Remember that Japanese /r/ can be realized as [ɾ], yet it has various other allophones, including [l].