Perceptual Distinction of English /l/ and /ɹ/ by Japanese Native Speakers

A Phonetic Study and Analysis

Author: Phoebe Lam, Kristine Singson and Sean Chan

Keywords: Phonetic Perception, Second Language Acquisition, Speech Production

Abstract:

The conventional belief of language acquisition studies assumes that language learners must first be able to perceive phonetic differences between sounds before they can produce said sounds distinctly. In particular, second language acquisition theorists typically assume that phonologically distinct sound production in L2 (the second language) cannot happen without its perceptual distinction. However, existing research regarding native Japanese speakers learning English challenges this framework. While we were unable to conduct a full study due to constraints of time and resources, we were interested in investigating sound perception in Japanese speakers with a low level of proficiency in English. In this study, we tested five participants who were native Japanese speakers who had spent between six months to a year in Canada. Their ability to differentiate /l/ and /ɹ/ sounds in minimal pairs were tested and graded, with the results analyzed for patterns or tendencies.

Background

In language acquisition studies, it is often thought that phonological errors made during speech production in a second language learner fall into one of two categories: inability to perceive phonological distinction, or the learner can perceive the difference but cannot produce it. A study by Wrembel et al. (2022) demonstrates that second and third language learners score significantly better on phonological perception tasks than production tasks. These findings suggest that learners may often be able to perceive phonological distinctions while being unable to produce them accurately. However, research by Goto (1971) and Sheldon and Strange (1982) contradict this framework and suggests instead that Japanese speakers learning English may be able to produce certain phonological differences even without perceiving it. This prompts investigation regarding the applicability of the typical production error categorization on Japanese learners of English. Indeed, it may be the case that this group of English learners produce phonetic differences without being able to accurately perceive their distinction. In an educational context, language instruction often emphasizes auditory feedback as a way for language learners to produce the "correct" phonemes. However, if production can, in fact, come before perception, this may not be the most effective method of teaching pronunciation. It is entirely possible that a focus on instructed articulation and proprioception may be more beneficial. Thus, exploring Canadian English speech sound perception and production in Japanese speakers may be crucial to second language acquisition and pedagogy.

The English sound system has a diversity of phonetic differentiations that do not exist in the Japanese language. One of the most salient examples among these is the Japanese use of the tap /ɾ/ compared to the English usage of /ɹ/ and /l/. The Japanese sound system uses the tap /ɾ/, romanized as "r", and does not contain the alveolar approximant /ɹ/ and alveolar lateral approximant /l/. The tap /ɾ/ is not typically spelled with an "r" in Canadian English. This sound is sometimes described as a "soft d sound" and appears in words like "ladder" and "butter" in Canadian English, spelled with a "dd" or "tt". Conversely, English contains both the /ɹ/ and /l/ sound and they are phonologically distinct, appearing in words like "red" or "light" respectively. As a result, native Japanese speakers are unlikely to perceive and produce the /ɹ/ and /l/ sounds of Canadian English in the same way as a native English speaker. It has been suggested that native Japanese speakers with low levels of fluency often struggle to differentiate between these two sounds in production.

Research Questions

In this study, we are investigating if Japanese native speakers can perceive the distinction of /ɹ/ and /l/ in Canadian English. We hope to gain better insight into the level of perception that a native Japanese speaker might have between these sounds, which are not phonologically distinct in their native language. As well, the study will aim to explore whether /ɹ/ and /l/ are easier to perceive in the onset or following a consonant, and if this differs depending on the vowel that occurs after the target sound. By testing a variety of phonological environments that /ɹ/ and /l/ occur in, we hope to find potential patterns in the ease of recognition based on the environment. As such, our two central exploratory research questions are as follows: if Japanese speakers can distinguish the difference between /ɹ/ and /l/ sounds, and if the phonological environment impacts the ability to differentiate.

Methodology

The experiment used 12 minimal pairs for a total of 24 words. Minimal pairs are defined as word pairs that are identical in pronunciation with the exception of one phoneme, which can then be used to isolate and test those specific sounds. Of the words used in this experiment, 8 pairs contained the target sounds /ɹ/ and /l/ while the remaining 4 contained /b/ and /v/ differentiation and

served as distractors. Within the 8 target minimal pairs, 4 of them contain words that start with the target sound, such as “light” and “right”, while the other 4 contain words which use the sound in consonant clusters, such as “fly” and “fry”. However, there were no words which ended in the target sound as minimal pairs that fulfill this distribution do not exist in English. There were some near minimal pairs, such as “car” and "call", but we decided against the inclusion of near minimal pairs as we could not guarantee possible influence of any phonological rules in these cases.

The word list is as follows:

1. Read
2. Fly
3. Berry
4. Clock
5. Root
6. Vowel
7. Gland
8. Right
9. Bat
10. Pray
11. Rip
12. Bowel
13. Loot
14. Vain
15. Grand
16. Lead
17. Play
18. Very
19. Light
20. Bane
21. Crock
22. Lip
23. Fry
24. Vat

The study was conducted through a facilitator, who met with several Japanese youths, either in person or over zoom, one at a time. Directions were given in both Japanese and English. The participants were given a fill-in-the-blank questionnaire where the target consonants were left blank, and with the rest of the word provided for them. They were then asked to listen to a single continuous audio file of a native English speaker reading the words aloud in order, with a brief pause between each word. See below for an example of the test provided:

> 音声を一回聞いて、直感で空欄を埋めてください！同じ単語が何回も出て来る可能性があります。お願いします。

Please play the audio once and fill in the blank based on what you hear, trust your intuition!! It is possible that certain words may appear multiple times.

1. *_ead*
2. *F_y*
3. *_erry*
4. *C_ock*
5. *_oot*
6. *_owel*
7. *G_and*
8. *_ight*
9. *_at*
10. *P_ay*
11. *_ ip*
12. *_owel*
13. *_oot*
14. *_ain*
15. *G_and*
16. *_ead*
17. *P_ay*
18. *_ery*
19. *_ight*
20. *_ane*
21. *C_ock*
22. *_ip*
23. *F_y*
24. *_at*

A pilot study was conducted to ensure the effectiveness of the experimental design and materials. A Japanese exchange student was asked to perform the task and provide feedback to the researchers and her results were analyzed for any possible patterns that could have been a result of experimental design. After completing the task, the participant expressed that she found the audio too fast and struggled to determine and write down the correct sound in the time allotted. This was fixed in the experiment, a new audio file was recorded with a 2-3 second pause between each word. Further, the facilitator noted that the participant expressed desire to listen again and was inclined to search words up on the internet to confirm the spelling. As a result, it was determined that it is preferable for a facilitator to be present for each participant to guarantee experimental integrity and consistency.

Participants

The study was conducted with a total of 5 participants, with ages ranging from 20 to 24. Each individual had spent between 6 months to a year living in Canada, learned English since middle school, and continued to study the language at the university level. Unfortunately, due to the limitations of time and resources, 4 out of 5 trials were conducted over video chat. Detailed participant information can be found in Figure 1.

Participant	Age	English Education	Duration in Canada	Region of Origin	Familiarity with keyboard
Participant #1	21	Elementary (1hr per week), Middle school, High school and UofA EAP 103 and EAP 135	7 months	Niigata	yes
Participant #2	24	Middle school, 3 years high school and 2 university classes at 北海道教育大学岩見沢校	1 year	Hokkaido	yes
Participant #3	20	3 middle school, 3 years High school and self study 1 hour a day	1 year	Shiga	yes
Participant #4	22	Middle school to university	8 months	Kyoto	yes but "bad" in typing English (prefers handwritten)
Participant #5	20	Middle school to university	6 months	Niigata	Yes but prefers handwritten

Figure 1: Participant Information Chart.

Results and Analysis

The results can be found in Figure 2. The table is given in the same order as that which the participants were given the words in. However, all distractors have been removed for clarity. In the table, green is used to denote correct responses and red for incorrect ones. The far-right column totals all five responses; here, green is used for words that received a score of 4 or more, red used for words with 1 or less, and yellow for everything else. The participants' individual totals are provided at the bottom of the table.

As seen, all participants correctly identified the words "read" and "fly", while none were successful in identifying "pray" or "gland". The results do not appear to be influenced by the type of vowel nor whether or not the word contained a consonant cluster. In the aforementioned cases, "read" and "gland" are followed by front vowels, whereas "fly" and "pray" are followed by diphthongs. In one notable case, a participant responded with "w" for "rip", which was an unexpected response. The experiment was designed under the assumption that participants would only struggle between alveolar sounds and did not account for any other type of misidentification. Interestingly, one participant requested to repeat the task, which was allowed. Their second set of responses differed from the first, yet did not result in improved accuracy. This data was excluded from the analysis. Figure 3 further divides up the score, with each participant's individual performance calculated, and a total at the bottom. Experimental score is defined as the number of correct responses out of all the times that "L" or "R" should have appeared, whereas accuracy scores are calculated out of the number of times "L" and "R" were given as a response. Hence, all the experimental scores were out of eight, while the accuracy score varied.

L/R Words	Participant #1	Participant #2	Participant #3	Participant #4	Participant #5	Overall
1. Read	R	R	R	R	R	5/5
0. Fly	L	L	L	L	L	5/5
0. Clock	L	L	R	L	L	4/5
0. Root	L	L	R	R	R	3/5
0. Gland	R	R	R	R	R	0/5
0. Right	R	L	R	L	R	3/5
0. Pray	L	L	L	L	L	0/5
0. Rip	L	W	R	L	L	1/5
0. Loot	L	R	R	R	R	1/5
0. Grand	R	L	R	R	L	3/5
0. Lead	L	R	L	L	R	3/5
0. Play	L	L	L	L	R	4/5
0. Light	L	R	L	L	R	3/5
0. Crock	L	R	L	L	L	1/5
0. Lip	L	R	L	L	L	4/5
0. Fry	L	R	L	L	L	1/5
Total	10/16	6/16	10/16	10/16	6/16	42/80

Figure 2: Participant Results.

Participant #1 Raw: 18/24 (75%)		L	R	Total
	Experimental Score	7/8 (88%)	3/8 (38%)	**10/16 (63%)**
	Accuracy Score	7/12 (58%)	3/4 (75%)	10/16 (63%)
Participant #2 Raw: 12/24 (50%)		L	R	Total
	Experimental Score	3/8 (38%)	3/8 (38%)	**6/16 (38%)**
	Accuracy Score	3/7 (43%)	3/8 (38%)	6/15 (40%)
Participant #3 Raw: 16/24 (67%)		L	R	Total
	Experimental Score	5/8 (63%)	5/8 (63%)	**10/16 (63%)**
	Accuracy Score	5/8 (63%)	5/8 (63%)	10/16 (63%)
Participant #4 Raw: 11/24 (46%)		L	R	Total
	Experimental Score	3/8 (38%)	3/8 (38%)	**6/16 (38%)**
	Accuracy Score	3/8 (38%)	3/8 (38%)	6/16 (38%)
Participant #5 Raw: 13/24 (54%)		L	R	Total
	Experimental Score	6/8 (75%)	4/8 (50%)	**10/16 (63%)**
	Accuracy Score	6/11(55%)	4/7 (57%)	10/18 (56%)
Total Raw: 70/120 (57%)		L	R	Total
	Experimental Score	24/40 (60%)	18/40 (45%)	**42/80 (53%)**
	Accuracy Score	24/46 (52%)	18/35 (51%)	42/81 (52%)

Figure 3: Expanded Participant Results.

Overall, the participants performed slightly better at identifying /l/ sounds compared to /ɹ/ sounds, though the difference was not substantial. This may be because they tended to choose /l/ more frequently overall. While /l/ sounds have received a test score of 24/40 between all 5 participants, the accuracy score per use is actually 24/46. In other words, the higher percentage score is a result of "L" being selected more often, and is not necessarily a result of a higher ability to perceive the sound. The reverse is also true for the /ɹ/ sound– although the sound received a test score of 17/40, the accuracy score was 18/35. The level of accuracy was at chance and seems to vary without a distinct pattern or influence. In examining further categories, similar results were found. With the score divided by vowel type, the totaled performance of all participants was 16/30 on front vowels, 9/20 on back vowels, and 16/30 on diphthongs. Participants' accuracy based on the placement of the target sound within the word was 23/40 for those at the beginning of the word and 18/40 in consonant clusters. While there was a higher success rate in the words starting with the target sound, it was not by a significant margin and the evidence is inconclusive. As such, the phonological environment did not seem to impact the level of accuracy and no clear patterns can be pointed out. Overall, the results show some tendencies, but it's difficult to draw definitive conclusions due to the small sample size of 5 participants. Given the relatively minor differences observed in the results, it seemed that the performance was at chance level.

Discussion

Based on the results of the experiment, it can be seen that Native Japanese speakers have significant difficulties distinguishing the difference between /l/ and /ɹ/. While we initially hypothesized that the phonological environment of the word was a significant factor, that also did not seem to be the case. However, we have some theories regarding the possible factors and limitations that could have influenced the results.

Potential Influences

1. **Order Effect:** Participants may have been influenced by the order in which sounds appear as all participants correctly identified the first two experimental items.

2. **Cognitive Load:** Hearing a high frequency of /l/ and /ɹ/ sounds in a row may have affected participants' sound perception. This, too, may have contributed to all participants correctly identifying the first two items.

3. **Word Familiarity:** While it seemed unlikely that participants avoided unknown words, this possibility cannot be entirely ruled out. When asked, participants claimed that they knew some of the words they got wrong. As well, one of the participants expressed uncertainty regarding how "real" the words were. Thus, we believe that it is unlikely that word familiarity played a factor.

Methodological Limitations

1. **Participant Constraints:** Time and participant availability were significant limitations. Initially, an in-person task would have been preferred, however this proved difficult. We had to recruit participants online as a result, but this introduced unexpected challenges. Each participant listened to the audio with a different device, leading to inconsistency with the audio quality. As well, it introduced the need for confirmation of each participants' keyboard competency and computer literacy to ensure that conducting the experiment online did not take away focus on sound perception.

2. **Demographic:** The small number of participants made it difficult to draw strong conclusions. A larger group of participants would have increased the reliability and validity of the results, especially if similar trends had persisted.

3. **Stimulus Design:** The lists were not counterbalanced across participants, which could have minimized the order effect.

    a. The use of capitalization in the fill-in-the-blank task may have introduced inconsistencies; using all lowercase letters would have been preferable.

    b. More distractor items could have been included to better isolate experimental data as the current list consisting of 2/3 experimental items is disproportionately high.

4. **Task Length:** Efforts were made to keep the experiment brief to avoid participant fatigue, however, this limited the scope of the data collected.

Conclusion

Our task was to explore if Japanese speakers can distinguish the /l/ and /ɹ/ sounds, as well as investigating any influence from the phonological environment in which these sounds appear. Although the conditions of our study were not ideal and there were confounds that should not be disregarded, we believe that our results hold valid. Our participants were unable to distinguish the two sounds, regardless of phonological environment and the scores were at chance across all categories. As also demonstrated in a study by Miyawaki et al (2022), Japanese speakers score significantly worse in speech sound perception compared to their English speaking counterparts. However, the aforementioned study showed that Japanese participants score above chance level. As such, we believe that a more thorough study, with a greater focus on phonological environments may yield more conclusive results.

A nuanced understanding of how phonological environments impact speech sound perception would be highly beneficial to understanding second language acquisition and language pedagogy. The conventional belief regarding order of acquisition may lack nuance and provide a one dimensional approach to language teaching. As such, we believe that further research should be conducted in order to understand the subtlety of sound perception by Japanese speakers.

Works Cited

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds "L" and "R." *Neuropsychologia, 9*(3), 317–323. https://doi.org/10.1016/0028-3932(71)90027-3.

Miyawaki, K., Jenkins, J.J., Strange, W., et al. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics* 18, 331–340. https://doi.org/10.3758/BF03211209.

Wrembel, M., Gut, U., Kopečková, R., & Balas, A. (2022). The relationship between the perception and production of L2 and L3 rhotics in young multilinguals; an exploratory cross-linguistic study. *International Journal of Multilingualism*, *21*(1), 92–111.

Sheldon, A., & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics*, *3*(3), 243–261.

Tyler, M.D. (2021). Perceived Phonological Overlap in Second-Language Categories: The Acquisition of English /r/ and /l/ by Japanese Native Listeners. *Languages*, 6(1), 4.