Department of Languages and Literature

Perception of English /l/ and /r/ by Japanese listeners - the influence of living abroad

Axel Hooper

Supervisor: Yasuko Nagano-Madsen

Examiner: Lars Larm

Abstract

Graduation thesis:	15 hp
Program and/or course:	Japanska, fördjupningskurs I avseende kandidatexamen
Level:	Undergraduate
Term/Year:	HT/2018
Supervisor:	Yasuko Nagano-Madsen
Examinator:	Lars Larm
Report Number:	xx (ifylles ej av studenten/studenterna
Keywords:	Phoneme, perception, allophone

Purpose: This thesis explores the topic of perception of /l/ and /r/ phonemes in Japanese speakers, primarily divided into two groups: Speakers who have lived in Sweden for 3 or more years and speak conversational Swedish, and average Japanese speakers, in an attempt to observe the effects learning Swedish has on one's English.

Theory: A hypothesis dubbed the Speech Learning Model (SLM) as well as my own hypotheses were used as a base for comparison. I primarily predicted an improved perception of /l/ relative to /r/ among the former group, as well as better overall perception.

Method: A form was created and distributed, using 15 minimal pairs to test perception among the aforementioned groups.

Results: Results showed consistently better overall perception among those proficient in Swedish compared to other Japanese speakers, though neither phoneme appeared to be significantly easier to perceive among either group, despite the similarity of /l/ and dissimilarity of /r/ phonemes between English and Swedish.

Acknowledgements

I would primarily like to thank Yasuko Nagano-Madsen, for helping me find a topic, introducing me to and allowing me to borrow various literature, as well as being my supervisor, providing help and consultation whenever it was needed, and being a great teacher. Without her, this thesis would not have existed. I would also like to thank my mother (Jill Hooper) for helping me with data analysis and Microsoft Excel, saving me considerable time and allowing me to easily create charts for use in this thesis. Another thank you goes to Lars Larm, for providing me with useful advice following my presentation, taking the time to read and grade this thesis, and being an all-around fantastic teacher. Finally, I would like to thank all who have responded to my form (including friends), and my friends and family who have supported me in writing, particularly Oliver Egan and Andreas Wahlberg for reading my thesis and providing feedback in its early stages.

Table of Contents

1 Introduction 1
1.1 Background 1
1.2 Problem, aim, and research questions. 1
2.Theories on L2 sound learning 3
2.1 Difficulties with foreign sounds in general 3
3. /l/ and /r/ in Japanese, English, and Swedish 4
3.1 /r/ in Japanese 4
3.1.1 History and dialectal variations 4
3.1.2 Japanese /r/ as an L2 sound 5
3.1.3 Allophones of /r/ in Japanese 6
3.2 /l/ and /r/ English 8
3.3 /l/ and /r/ in Swedish 9
3.4 Summary of phonetic differences 10
4. Previous studies on the perception of English /l/ and /r/ by Japanese listeners 11
5. Perception test 13
5.1 Wordlist and stimuli 13
5.2 Subjects 15
5.3 Questionnaire 15
5.4 Procedure 16
5.5 Hypotheses and prediction 16
6. Results 19
6.1 Overall identification by the two groups 19
6.2 Comparison by position in a word (word type) 21
6.3 Comparison of /l/ and /r/ 22
6.3.1 Phoneme identification by the two groups 22
6.4 Comparison by gender and background 25
6.5 Comparison with Shimizu and Dantsuji's results 27

7. Discussion 28
7.1 Research Question #1 28
7.2 Research Question #2 28
7.3 Research Question #3 30
7.4 Research Question #4 30
7.5 Research Question #5 30
7.6 Further research 31
8. Conclusion 33
9. References 34
10. Appendix 37

1 Introduction

1.1 Background

This thesis concerns the topic of perception of the English consonant phonemes /l/ and /r/ by native speakers of Japanese who are long-term residents of Sweden. While English and Japanese have a considerable number of differing qualities in pronunciation, one of the most challenging of which for Japanese speakers is arguably the differentiation of the English /l/ and /r/, both in production and perception. Difficulties in differentiating these two sounds among Japanese speakers has been well-documented for a long time (Goto 1971, Miyawaki et al., 1975), and many studies have been performed (Shimizu and Dantsuji 1983; Bradlow et al. 1999, Aoyama et al. 2004) on Japanese speakers in order to test quality of perception as well as production of these two sounds, in an effort to pinpoint perceived difficulty and/or train learners to better differentiate them. Most of these studies (such as Aoyama, K et al. (2004)) examine either Japanese natives living in Japan, or those living in English-speaking countries long-term. Some studies also examine young children, to observe the effects age has on perception (Aoyama et al. 2004). Those examining adults (Goto 1971, Miyawaki et al. 1975, Aoyama et al. 2004) are all unclear as to whether or not they have examined adults not of typical student age (early 20's).

1.2 Problem, aim, and research questions.

Seemingly few to none of said studies have examined the effects of Japanese nationals living long-term outside of both Japan and the Anglosphere. This creates a gap of sorts; the targets of these studies are all either in an English-speaking environment, or they are completely immersed within their own language, one which does not possess the luxury of distinction between /l/ and /r/. They are, simply put, either in or out, not in between. The aim of the

present study is to examine if the experience of living abroad (here in Sweden) can influence the perception of the English /l/ and /r/ by Japanese listeners. In an attempt to uncover new knowledge, I have elected to fill the aforementioned gap, and thereby pose and determine the answer to the following research questions:

- Does living in Sweden and being exposed to Swedish, which also has a distinction between the /l/ and /r/ phonemes, improve the perception of /l/ and /r/ in English by Japanese natives?
- Is there a difference in identification depending on the position of a phoneme in a given word, e.g initially (single vs. cluster) vs. medially (single)?
- Among the aforementioned group, is perception of the English /l/, arguably equal to the Swedish equivalent, greater than that of /r/?
- Are there any differences relating to the subject's gender and their background, such as English Major influence?
- To what extent can results of the present study contribute to the existing theories of acquiring /l/ and /r/ phonemes in English by Japanese speakers?

2.Theories on L2 sound learning

2.1 Difficulties with foreign sounds in general

Many individuals can relate to having trouble pronouncing sounds in foreign languages. One could argue that the reason Japanese speakers appear to experience considerable difficulty with /l/ and /r/ in English is that such a distinction does not exist in Japanese phonology. However, McMahon (2002) brings up the fact that one's native phoneme system can get in the way of learning a foreign language. More specifically, she states that sounds that are phonemically contrastive in the target language, yet allophones of a single phoneme in one's native tongue, can be particularly difficult to learn, while new phonemes entirely unknown to the learner are easier due to unfamiliarity. She goes on to say that an unfamiliar sound must be learned from scratch, whereas an old sound in a new role requires suppression of one's instinctive pronunciation habits. Consequently, she applies this to perception too; a new sound is simple to identify due to unfamiliarity, whereas differentiating two allophones one has previously thought of as the same is difficult. This idea is backed up by Flege (1995), in which he hypothesizes: "The greater the perceived phonetic dissimilarity between an L2 sound and the closest L1 sound, the more likely it is that phonetic differences between the sounds will be discerned", in which L1 is one's native tongue and L2 is the target language. Flege appears to have dubbed this, among other hypotheses as well as postulates, as the "Speech Learning Model" (SLM), and this concept shall therefore be referred to as such henceforth.

3. /l/ and /r/ in Japanese, English, and Swedish

In order to make clear why Japanese people may find /l/ and /r/ so difficult, one must first define what these sounds are, and in what shape they take form in in each respective language. Hence, this section will discuss the phonetic qualities of /l/ and/or /r/ in languages relevant to this study.

3.1 /r/ in Japanese

3.1.1 History and dialectal variations

In ancient Japanese, there were no words that begin with an /r/ phoneme. In Modern Japanese, any words that begin with an /r/ phoneme are said to belong to one of the following categories:

- Foreign words such as *gairaigo* (words of foreign and non-Chinese origin), *kango* (words of Chinese origin), and words of ainu origin.
- Auxiliary words and particles (“rareru”, “rashii”, and so on)
- A small number of onomatopoeic words such as “ranran”, or “runrun”.

In short, the vast majority of words that start with an /r/ are not of Japanese origin.

The dialectal variation or /r/ has also been known. Sugito (1984) writes as follows:

“The za-gyou da-gyou, and ra-gyou in Japanese have many similarities, and “da, de, do” getting mixed up with “ra, re, ro” is not uncommon. The presence of confusion with these sounds has been seen in Edo period texts, and is thought to be fairly old. In western Japan, particularly the Kansai region, these sounds are often mixed up, but this trend is seen to a lesser extent all over Japan”. (My translation)

It should be also added that the /r/ phoneme has an allophone as a trill [r] in Tokyo downtown speech, which is known as *makijita*「巻き舌」in a *beranmee* style「べらんめえ口調」. (まき- じた【巻(き)舌】) (discussed later).

3.1.2 Japanese /r/ as an L2 sound

The difficulty in acquiring the phonetic realisation of Japanese /r/ by learners of Japanese has been documented for some time. In the booklet *nihongo no tokushoku*「日本語の特色」(1979:88), a good summary can be found. For Westerners, one of the difficulties is the consonant in the *ra-gyou* (ラ行の子音). In particular, Americans pronounce the Japanese [d] sound with a weak plosive so that /edobungaku/「江戸文学」(Edo period literature)」sounds like [ero bungaku]「エロ文学」(=erotic literature). It also points out that among Chinese speakers, Cantonese speakers tend to mix up /na/ and /ra/ in Japanese.

Sugito (1984):

"These consonants are difficult to differentiate for foreign students of Japanese, and according to the results of a perception test consisting of meaningless words such as "doza", "rezo", and "roda" performed on Americans with pronunciation by speakers of Tokyo Japanese, it was found that discrimination between the /r/ and /d/ consonants was particularly difficult." (My translation)

Difficulty in producing the Japanese /r/ by American learners is discussed in Aoki (1990:228) in more detail. Aoki concludes that it is not so problematic for American English speakers to learn Japanese consonants in general that will lead to confusion in meaning since the number of consonant phonemes in Japanese is much less than that in English. This is true for /r/ since there is no contrast between /l/ and /r/ in Japanese. However, he does mention that there is a

considerable difference in the phonetic realisations of /r/ produced by American learners. He writes: “the English /r/ is an ‘approximant’, in which the tongue never touches the hard palate, Japanese /r/ does (whichever allophone it is). As a result, a syllable with the American /r/ (an approximant) becomes longer”. (My translation)

Swedish learners are mentioned in Nagano-Madsen (2005) as well as in (Nagano-Madsen et al. 2014). In the latter, the realisation of /r/ in Japanese is difficult for Swedish learners among beginners, but it is improved drastically for those who have studied in Japan. Replacing the Japanese /r/ with the Swedish /r/ gives an impression of ‘foreign accent’ but it does not result in phonological confusion. This was in contrast with some other phonetic features such as manifestation of long vs. short vowels and consonants, which was persistently difficult even after spending a year in Japan.

3.1.3 Allophones of /r/ in Japanese

In a relatively recent account proposed by Vance (2008), /r/ in Japanese is described as possessing only one liquid phoneme, typically pronounced as an alveolar tap [ɾ]. He likens this to the American pronunciation of the words *patty* and *paddy* [pʰæɾi], in the event that they are pronounced identically. Vance goes on to briefly describe the presence of the voiced apico-alveolar trill [r], which he claims is stereotypically associated with gangsters, but marginal overall within the linguistic spectrum of Japanese due to many speakers having difficulty with its pronunciation. McMahon (2002) notes that this sound is uncommon for speakers of English. We will discuss this sound further later on.

However, in books written by Japanese phoneticians, one can find more detailed accounts, which do not show a total agreement. They differ in the exact phonetic realizations of the /r/ phoneme as well as the environment in which a given allophone appears.

Amanuma et al. (1989) describe the consonant phonemes in the Japanese ラ行 (*ra-gyou*, the five characters containing the Japanese "r") as having three manners of articulation. These are the lateral, trill, and the tap.

Lateral [l]

In the phrase "あっら、まあ。" (arra, maa "Oh my, ah."), the "ra" is described as being pronounced similarly to [l]. They also describe pronouncing the *ra-gyou* ラ行 (=ra-row) as [l] at the word initial position is common, and that there are individuals who always pronounce the ra-gyou as [l] regardless of position. This points to [l] possibly being an existing phoneme in Japanese.

Trill [r]

Regarding the trill, they show an illustration of a tongue making contact with the alveolar ridge, and say that this sound is produced by using the tip of the tongue to produce a flap motion multiple times. This is described as being mostly unused in normal Japanese, however they also liken it to the sound used when speaking with *makijita* 巻き舌 (=a Japanese layman term for trilled /r/ phonemes).

Tap [ɾ]

When discussing the trill [r], they claim that performing the flap motion once instead of repeatedly creates the alveolar tap [ɾ]. The latter is portrayed as the sound that appears word medially, using examples where many pronounce the word initial /r/ phoneme as [l], and subsequent ones as [ɾ] (e.g りょうり, *ryouri*, "cuisine").

In short, Amanuma et al. present standard Japanese as mainly having the [l] and [ɾ] allophones for /r/, with speaker variation being an important factor in articulation.

Like Amanuma et al.(1989), Aoki (1990) also claims that the Japanese /r/ phoneme has three major allophones. However, the two studies differ as follows. Word initially, Aoki (1990:228) proposes “a kind of plosive with weak plosion, which is produced by a weak closure made by the tongue tip/back and alveolar ridge” (My translation), while Amanuma et al. postulate [l] in this position. Word medially, Aoki also postulates *hajiki-on*「弾き音」(=flap). However, from the description, it can be interpreted as though Aoki also postulates a tap. Aoki also includes [l] as an allophone of /r/, noting the lateral [l] as mainly appearing in singing.

Kitahara et al. (2003) describe the consonants of the ra-gyou as consisting of the alveolar tap [ɾ] in which the tip of the tongue, narrowly and lightly coming into contact with the alveolar ridge for a short period of time, creates a stop. He goes on to describe the lateral tap [ɺ] as also existing according to other sources.

Interestingly, Like Aoki (1989), Kitahara et al. describe the Japanese /r/ as possessing similarity towards [d]. Note, however, Aoki suggests [d] only in the word initial position while Kitahara et al. doesn’t.. The main difference between [d] and [ɾ], according to them, is described as being the area of contact. For [d], it is wide whereas in the case of [ɾ] it is narrow, and [d] also has a greater degree of plosivity. In words such as アッラー (arraa, “Allah”) or はんらん (*hanran*, “rebellion”), the location of the tongue is apparently the same as [ɾ], however the duration of contact towards the alveolar ridge is longer, and is more close to a plosive than a tap.

3.2 /l/ and /r/ English

McMahon claims that the voiced retroflex approximant [ɹ] is the most common realisation of /r/ for speakers of Southern British English and General American, however she goes on to note that different realisations of /r/ are found throughout the English-speaking world.

Regarding /l/, she defines it as a voiced alveolar lateral approximant, and lists it as one of four approximant consonant phonemes in English, and says that it has two main allophones: [l] and [ɬ].

3.3 /l/ and /r/ in Swedish

Riad (2013) defines the /r/-phoneme as an apical trill [r] in Central Swedish. He goes on to describe it as follows: “We characterize it phonologically as a retroflex coronal, where the retroflex property of the trill is taken to be central. The trill is pronounced in the alveolar region.”.

Regarding /l/, Riad claims that it has the dental [l] as its main allophone, while also deeming it a lateral approximant.

Elert (1966) describes the Swedish /l/ as being an example of a lateral consonant, and goes on to note that many articulate it in such a way that air flow passes by only one side of the tongue. Interestingly, in his list of retroflex consonants he lists [l̢] as being a common realisation of /l/ in uppsvenska (Swedish in Central as well as the north of Sweden) in words such as pärla (pearl) or sorla (murmur).

He states that the /r/ sound in uppsvenska is often a voiced alveolar trill [r]. Also of note is that in Stockholm in particular, a voiced alveolar fricative [z̢] is yet another realisation of /r/. While not specific to Swedish, much like McMahon he describes the English [l] as an alveolar consonant.

In a section on lateral phonemes, Engstrand (2004) defines /l/ phonemes in standard Swedish as being dental as a general rule of thumb, also noting that the IPA has the symbol [l̢] for a retroflex /l/, commonly used for words like porla (murmur).

3.4 Summary of phonetic differences

Riad clearly defines the Swedish /r/ as an apical trill [r], and this is seemingly in congruence with Elert as well as Engstrand, who all portray [r] as being common in standard Swedish (though with the presence of regional differences). In contrast, McMahon claims that: “Trilled [r] is now rather uncommon for speakers of English, although attempts at imitating Scots often involve furious rolling of [r]s”. This, as well as personal experience, leads me to believe that the English and Swedish /r/ are fundamentally different phonemes and thereby sounds in general. Conversely, Japanese as described by Vance and Amanuma et al., *does* possess the trilled [r], albeit exclusively when exaggerated, and some speakers are unable to pronounce it. The more common Japanese alveolar tap [ɾ], (according to Vance, Amanuma et al., Aoki, and Kitahara et al.) however, is of course different to that of English and Swedish variants of /r/. In English’s case, the much different [ɹ] is most prevalent, whereas with Swedish it is usually [r]. However, Japanese also has another variant of the sound which is more like [d] that appears word initially, or like [l] in singing as described in Aoki (1990). In short, one could generalise and say that English as uses the [ɹ], Swedish uses the [r], and Japanese uses the [ɾ].

With /l/, both English and Swedish appear to possess the sound, however it is alveolar in English according to McMahon, and dental in Swedish according to Riad and Engstrand. Both McMahon and Riad claim that it is a lateral approximant however, and so the minor difference of being dental vs alveolar, that is to say, different allophones, can be considered unimportant as it is a minor difference at best. Also worthy of note is that the Japanese alveolar trill [ɾ] is more perceptually similar to the English [l] than the English [ɹ] is, according to Takagi, as well as Sekiyama & Tohkura (cited in Aoyama K. et. Al. 2004). This is crucial in that, going by the Speech Learning Model, the English /l/ should therefore be more difficult for the Japanese to

acquire than /r/. Conversely, some sources, namely Aoki and Amanuma et al., claim that the /l/ actually exists in Japanese, albeit only in a limited form according to the former.

To summarise, the most common pronunciation of /r/ seems to differ a fair amount between all three languages, while /l/ is mostly the same in English and Swedish, and the existence of it in Japanese is debatable and will not be argued here as it is outside the scope of this thesis.

4. Previous studies on the perception of English /l/ and /r/ by Japanese listeners

Studies on /l/ and /r/ date back to 1970's, starting with that of Goto (1971), in which he describes the Japanese as possessing an intermediary sound between the English /l/ and /r/, and tests the perception as well as production of the two on Japanese and Americans. This was done by placing words containing these sounds in between those that do not, and asking respondents to identify them. A second experiment was performed using so-called *minimal pairs*, pairs of words in which only one differing phoneme is contained, and are otherwise phonetically identical.

Shimizu and Dantsuji (1983) performed a study in which they made use of minimal pairs in order to, like Goto, test perception. However, they also used some word pairs with more than one differing phoneme (e.g "secretly" vs. "secretary"), as well as combinations of letters of the alphabet that do not form actual words (e.g "ul" vs. "ur"), using these in the same manner as minimal pairs. They tested 32 undergraduate students of English literature. In their testing, they had three types of tests. Type 1 had respondents identify sounds in minimal pairs, in which one was pronounced after the other. Respondents were asked to identify the first word.

The second and third types (Type 2 and Type 3) consisted of “carrier sentences”, sentences in which the only variable would be the minimal pair word. The difference between the two types was whether the variable would be placed at the beginning or end of the sentence, however it was concluded that the location of the minimal pair word had no noticeable effect on perception.

Aoyama, K et al. (2004), claiming that “the English /r/ is perceptually more dissimilar from the Japanese /r/ than English /l/ for Native Japanese (NJ) speakers”, attempted to evaluate the Speech Learning Model established by Flege, by “investigating whether NJ speakers will have more success acquiring English /r/ than /l/”. In their experiment, they observed perception as well as production of /l/ and /r/ in both native Japanese children and adults living in the U.S. twice, with a one-year interval. They found that test subjects had shown more improvement in relation to /r/, rather than that of /l/, which supports the Speech Learning Model proposed by Flege.

5. Perception test

As established above, the goal of this study will be to attempt to find a relationship between long-term residence of Japanese people in Sweden, and the ability to clearly perceive the differences between the English /l/ and /r/ sounds. To do this, a survey was performed, testing auditory perception. The survey can be viewed at the following link: https://goo.gl/forms/mSp28f1RDRY7Qve53

5.1 Wordlist and stimuli

A total of 15 minimal pairs were prepared. They can be classified as follows:

1. Pairs in which [ɹ] or [l] are in the word initial position.
2. Pairs in which [ɹ] or [l] are in the word medial position.
3. Pairs in which [ɹ] or [l] are in the word initial position in a consonant cluster.

Words falling under each of these three categories will be referred to as “Type 1”, “Type 2”, and “Type 3” words respectively in the results section. Simple everyday words were chosen, and some of the words were adopted from Shimizu and Dantsuji (1983)’s study so that the results could be compared afterwards (see the “results” section).

Each word was recorded by the author of this thesis who is a native speaker of British English, and saved as a sound file. Each word was pronounced twice and the words were presented randomly in one continuous sound file.

The following is a list of the minimal pairs used as well as their categories into which they fall under. Categories are marked by a number in parentheses, and words are listed in the same order as the survey.

Categories:

“S&D” marks pairs appearing in Shimizu and Dantsuji’s study. (Word count: 7)

The number (1-3) marks the type. (Word count: 5 of each type, 15 total)

The letter (“R” or “L”) marks which word was pronounced in the survey (the correct answer) (Word count: 8 “R” answers, 7 “L” answers).

#1: Right/Light (S&D) (1) (L) word initial

#2: Pirate/Pilot (S&D) (2) (R) word initial

#3: Read/Lead (S&D) (1) (L) word initial

#4: Correct/Collect (2) (R) word medial

#5: Crime/Climb (3) (L) word initial, consonant cluster

#6: Raw/Law (S&D) (1) (R) word initial

#7: Pray/Play (S&D) (3) (R) word initial, consonant cluster

#8: Jerry/Jelly (2) (L) word medial, syllable initial

#9: Grass/Glass (3) (L) word initial in consonant cluster

#10: Rice/Lice (S&D) (1) (R) word initial

#11: Arrive/Alive (2) (L) word medial,

#12: Wrong/Long (S&D) (1) (R) word initial

#13: Berry/Belly (2) (L) word medial

#14: Bruise/Blues (3) (R) word initial, consonant cluster

#15: Fry/Fly (3) (R) word initial, consonant cluster

5.2 Subjects

In order to answer the questions posed, the target of this study was twofold; subjects were divided into two separate groups, and compared to each other. These are as follows:

Group 1: Native speakers of Japanese who have lived in Sweden for a minimum of three years, and possess at least conversational ability of Swedish.

Group 2: Native Japanese speakers that do not fulfil the above conditions.

With these two groups, the goal was to identify whether or not the former has greater perceptive ability than that of the latter.

5.3 Questionnaire

At the beginning of the survey, a short description about myself, as well as the contents of the survey was displayed, followed by a question in which respondents were asked if they have lived in Sweden for 3 years and speak conversational Swedish. This was done to identify which group (Group 1 or Group 2) they adhere to. Following this, respondents were separated into two separate sections consisting of the exact same layout. During this part of the survey, respondents were asked to provide their gender, as well as whether they have either previously or are currently majoring in English or not. Finally, a short text box describing how to fill out the survey preceded a sound file, followed then by a list of minimal pairs. Next to this list, the respondents had three options: “R”, “L”, or “聞き取れませんでした” (could not hear). In the minimal pair list, words containing /r/ phonemes were listed on the left, while those containing /l/ phonemes were listed on the right. Consequently, the “R” option was on the left, and the “L” option was on the right. This was done to ease the process of answering and reduce possible confusion. All of the words presented as well as their correct answers were ordered randomly in terms of type and correct answer, so as to not create

undesirable patterns within the results. The order of words in the survey is identical to the list provided above.

5.4 Procedure

The questionnaire was distributed digitally. Respondents would then go through the form, listening to each word one at a time, answering what they think they just heard. Once a sufficient amount of responses had been acquired, results were presented. Groups one and two were compared to each other, and words appearing in Shimizu and Dantsuji's study were compared to their results (although the methodology is different, it is similar enough to be of value).

Data from the form was compiled and then put into charts via Microsoft Excel. Said charts are the charts appearing later in the results and APPENDIX section.

5.5 Hypotheses and prediction

The Speech Learning Model established by Flege points to Japanese speakers being more likely to correctly identify the English [ɹ] than the [l]. One could however challenge this claim, and argue that living in Sweden whilst also learning Swedish would lead to a different result. The reasoning behind this is that Japanese living in Sweden are arguably exposed to [l] far more than the English [ɹ] due to the latter not existing in Swedish, leading to improved perception of the former. This is due to the Swedish and English [l] being the same phoneme, albeit different allophones of each other. It could be believed that respondents within Group 1 will have better overall perception of /l/ than that of /r/ in relation to Group 2. Given what the SLM says however, one could predict that the average native speaker of Japanese would likely perceive /r/ better than /l/, and so predict such an outcome within Group 2.

One could also think that Group 1 will possess better overall perception of both phonemes compared to Group 2. While the English [ɹ] does not exist in Swedish, exposure to Swedish could further solidify the English [ɹ] as a foreign sound, allowing the respondent to more clearly identify it as not being [l].

The perception of /l/ and /r/ phonemes may vary depending on the location of them in a word. A quick analysis (calculating the mean of correctly identified words of each type) of the true minimal pairs in Shimizu and Dantsuji's results provides the following data:

Percentage of Japanese subjects who correctly identified minimal pairs (mean):

Type 1 words: 86.89%

Type 2 words: 80.02%

Type 3 words: 67.95%

(Note: the number of words used in order of quantity divided by type goes 1>2>3. In other words, Type 1 words were tested the most by Shimizu and Dantsuji)

The above data points to subjects having more difficulty identifying both /l/ and /r/ the more "surrounded" the phoneme is. Type 1 words simply begin with either phoneme, meaning there is nothing previous to confuse or otherwise create difficulty for the listener.

Type 2 words have a vowel phoneme preceding the /r/ or /l/ phoneme, and the structure of Japanese is for the overwhelming majority of the time in sync with this: phoneme clusters are mostly either vowel-vowel or consonant-vowel; double consonants are comparatively few, with the only exceptions comprising of the Japanese characters ん (n) or っ (tsu).

Type 3 words have a consonant preceding the phoneme, and are arguably the most distant from Japanese in terms of general structure, making them seemingly the hardest to identify.

In their tests on carrier sentences, while the position of the pair (initially vs finally) was not deemed important, having a sentence for a given pair provided more difficulty than

presenting pairs in isolation. This matches the idea of phonemes that are “surrounded” being more difficult to identify.

For these reasons, one can expect the order of difficulty, and consequently, amount of words correctly identified, will be 1>2>3, in which each number corresponds to the types established above. This matches Shimizu and Dantsuji’s study, albeit on a larger scale. Those who have majored in English at university may have better perception than others among their group, simply as a result of greater exposure. What will be of interest is the degree of difference between those who have and those who have not. Finally, the present study also examines the effect of gender as an extra category.

6. Results

The form was distributed and was then open for responses for a period of 15 days. The total amount of respondents during this period was 60. More responses on both sides would of course have been ideal, however this was still enough to draw valuable conclusions. Also, having roughly 30 responses per group creates a similar amount to Shimizu and Dantsuji, meaning the comparison may be of greater overall quality than otherwise. Group 1 (Sweden) consisted of 32 individuals, exactly half (16) of which had majored in English. 27 (84.4%) were females, while the remaining 5 were male. Group 2 (Japan) consisted of 28 individuals, of which 13 were females and 15 (53.6%) were males, creating an overall more balanced pool of respondents in terms of gender. There were 16 English majors, and 12 who do/did not major in English.

6.1 Overall identification by the two groups

The results showing correct identification of /l/ and /r/ by the two groups are presented in Figure 1 below.

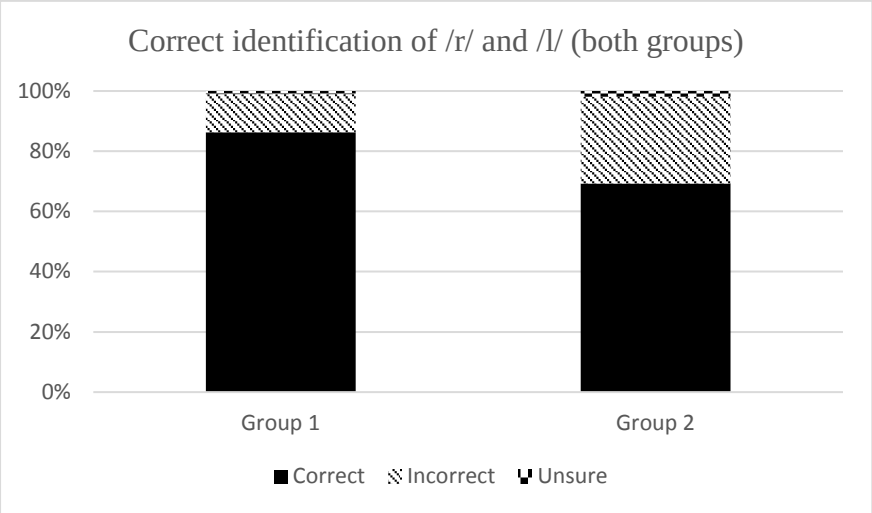

Figure 1: Correct and incorrect answers divided by group.

To examine the perception patterns by the two group in more detail, detailed information formation is presented in Figures 2 (a) and (b) below.

The pattern of perception in the two groups is very similar for some words. The top five highly-scored words for Group 1 (Sweden) are as follows: Correct (Q4) > Rice (Q10) > Pray (Q7) > Lead (Q3) > Jelly (Q8). Similarly, for Group 2, they are: Jelly > Rice/Alive (Q11) > Lead/Correct/Belly (Q13)

The least correctly identified words for Group 1 are: Climb (Q5) > Raw (Q6) > Bruise (Q14) > Pirate (Q2)/wrong (Q12). Group 2: Climb > Wrong > Pirate/Raw > Light (Q1). One can conclude that the perception patterns among both groups is similar, and that the difference is quantitative rather than qualitative.

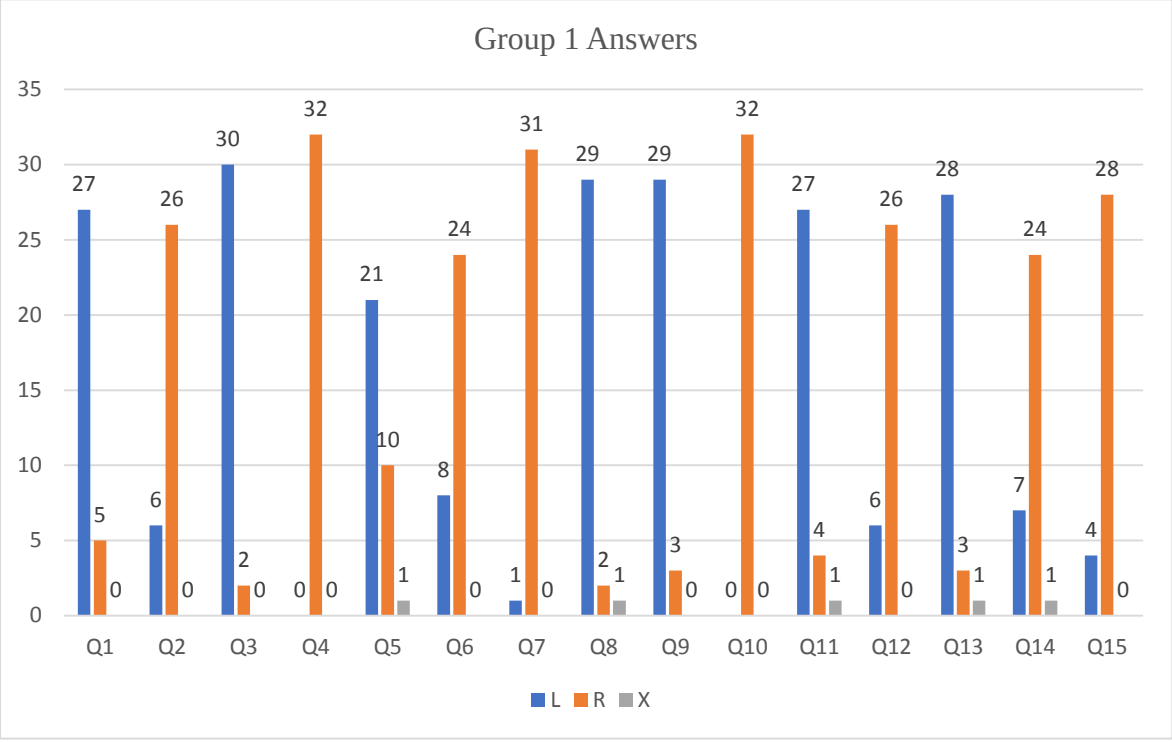

Figure 2 (a) Responses for individual words for Group 1 (Sweden).

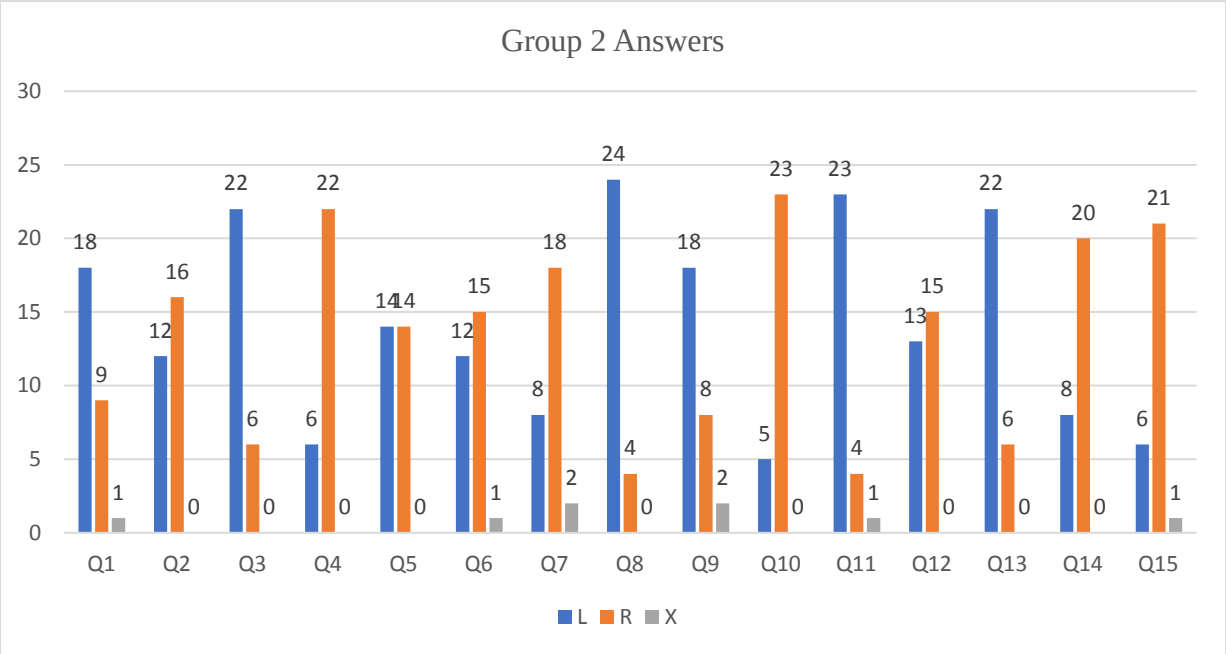

Figure 2 (b) Responses for individual words for Group 2 (Japan)

6.2 Comparison by position in a word (word type)

The results of the comparison by position in a word (initial vs. medial) as well as by structure (single vs. cluster) are presented in Figure 3 (a) and (b) below. They are referred to as word type 1 (word initial), word type 2 (word medial, single) and word type 3 (word initial in a consonant cluster). The correct scores are Type 2 > Type 1 >Type 3 for both groups. Both groups found the word medial identification easier and the consonant cluster most difficult. The difference in identification is larger for Group 2.

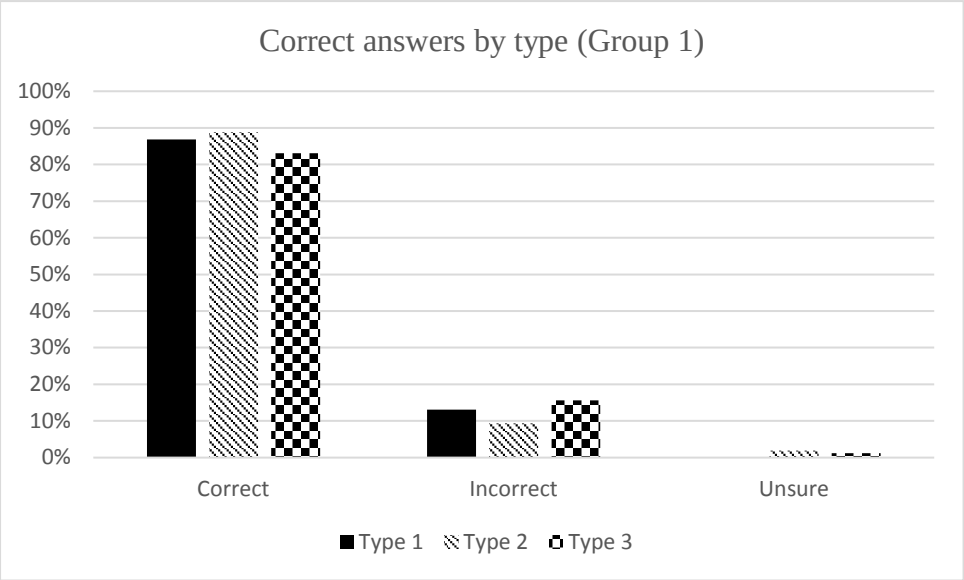

Figure 3 (a) Correct score by word type for Group 1 (Sweden)

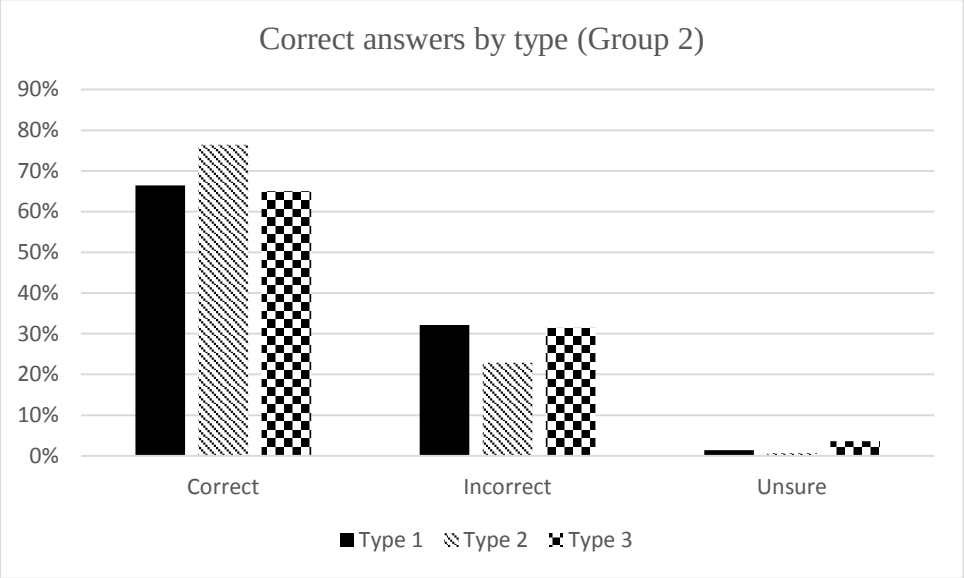

Figure 3 (b) Correct score by word type for Group 2 (Japan)

6.3 Comparison of /l/ and /r/

6.3.1 Phoneme identification by the two groups

Data was analysed and assembled into the following figure, representing correct answers based on phoneme, divided by group. Group 1 perceived /r/ better than /l/, and vice versa for Group 2. The gap was only 2 percentage points in the former, and 5 in the latter, however.

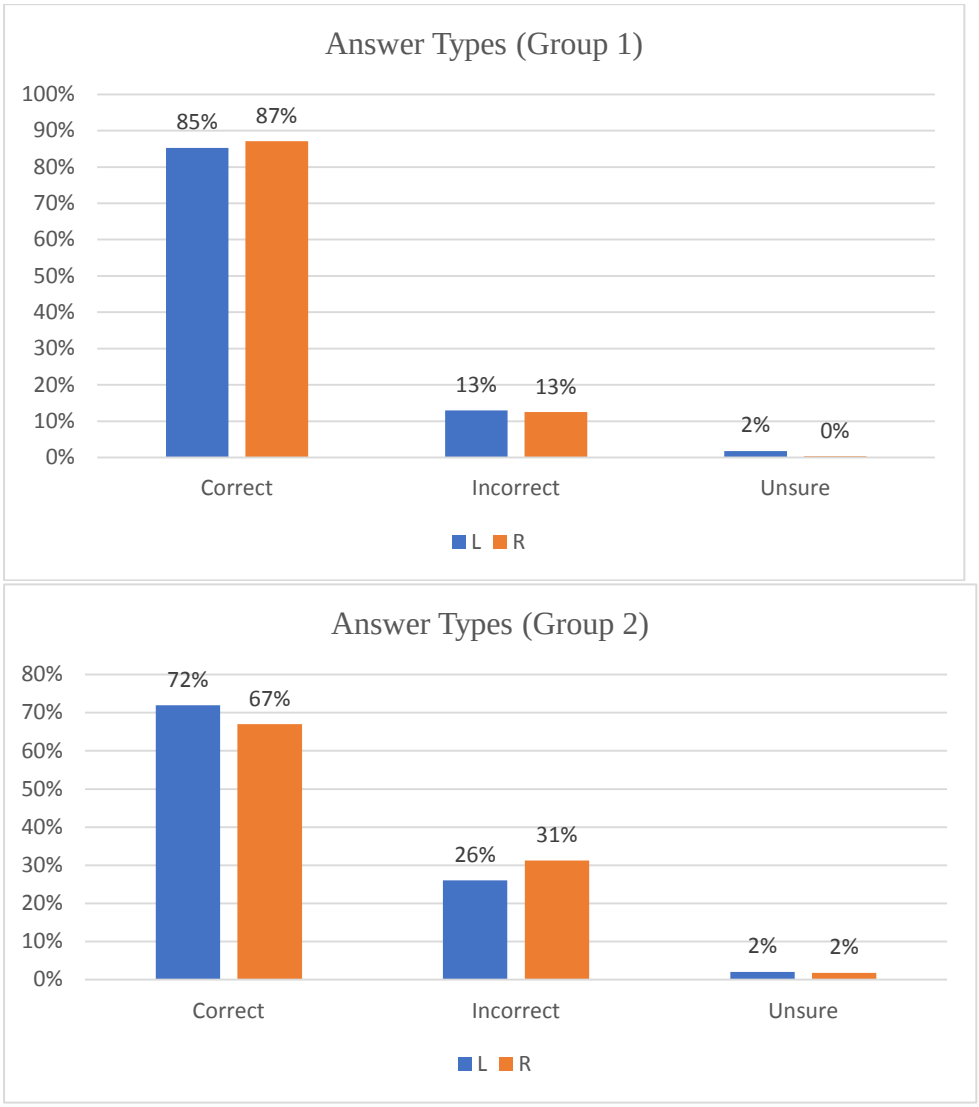

Figure 4: Summary of the perception of /l/ and /r/ for Group 1 (above) and Group 2 (below)

Further examination was conducted, in which /l/ and /r/ identification was divided based on word type. This data was gathered across both groups due to similar patterns. The next series of figures show how correctly /l/ and /r/ were perceived based on word type. Type 1 and 2

(initial and medial, single) words pronounced with /l/ were more correctly identified than /r/ by 6 percentage points, whereas type 3 (consonant cluster) words (the hardest type for both groups), had the opposite result, with /r/ being more correctly identified by 10 percentage points.

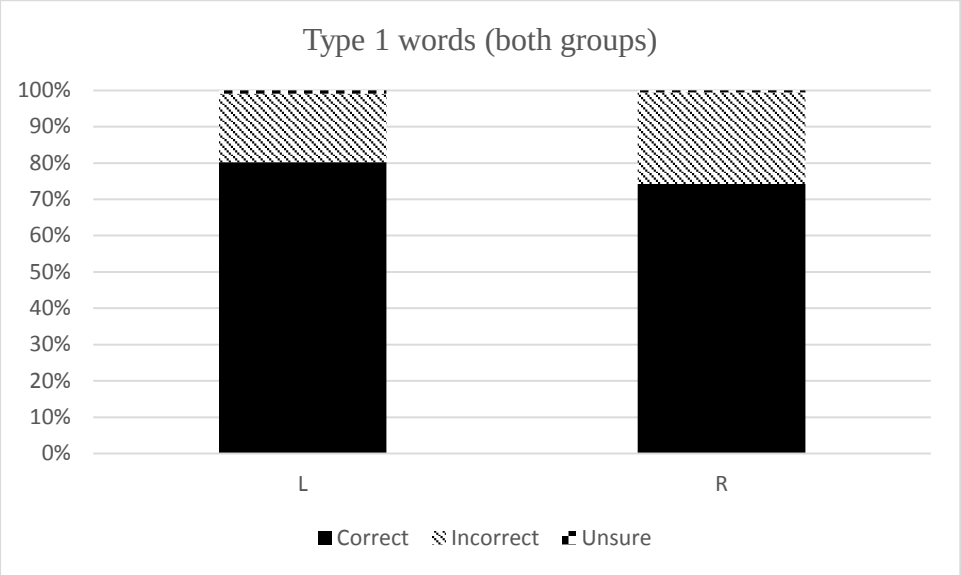

Figure 5 (a): Correct score of /r/ vs /l/ for Type 1 words. (Word initial).

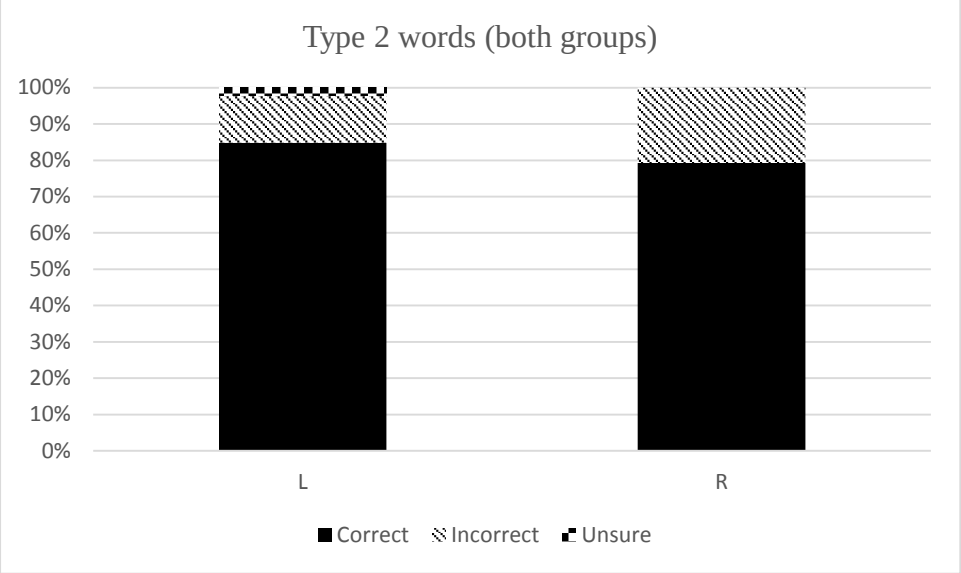

Figure 5 (b): Correct score of /r/ vs /l/ for Type 1 words. (Word medial).

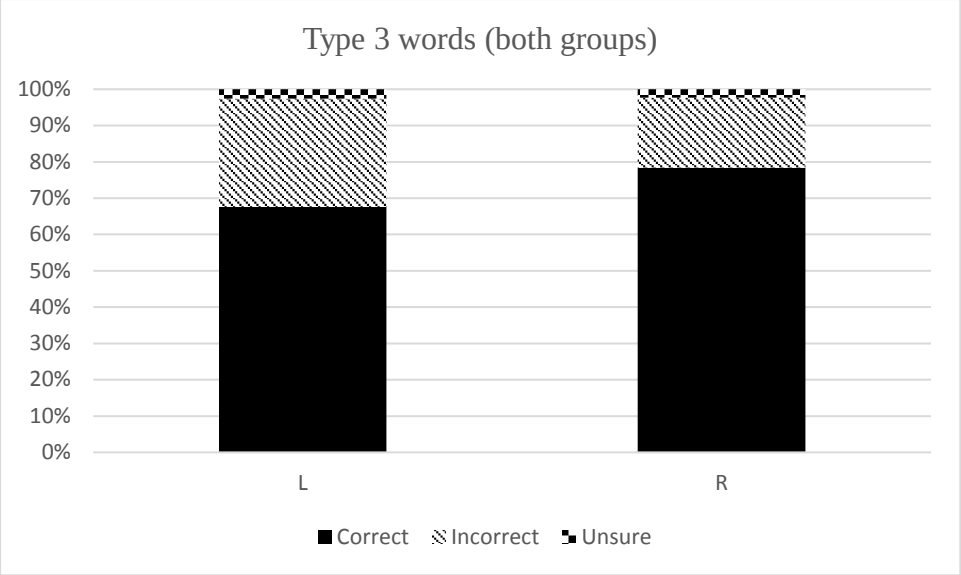

Figure 5 (c): Correct score of /r/ vs /l/ for Type 3 words (word initial in cluster).

6.4 Comparison by gender and background

On the whole, sex appeared to have some influence on responses (Cf. Figure 6 below). Both genders had fairly similar results in Group 1, though it should however be noted that the vast majority of respondents in said group were female, and that more male respondents or at least a higher percentage would create more reliable data than what is presented here. In group 2 however, females performed notably better than their male counterparts, with a whopping 13% disparity. Details for each group are presented in Figures 8-10 in APPENDIX.

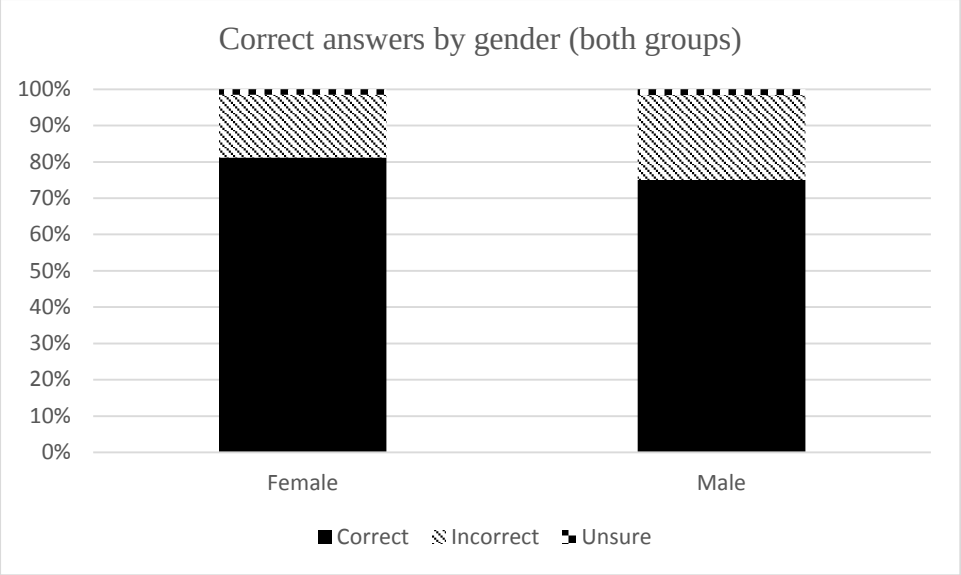

Figure 6: Gender differences

Interestingly, non-English Majors did better than their counterparts across the board, though not by a large amount. This indicates that having majored in English is not a deciding factor in terms of perception.

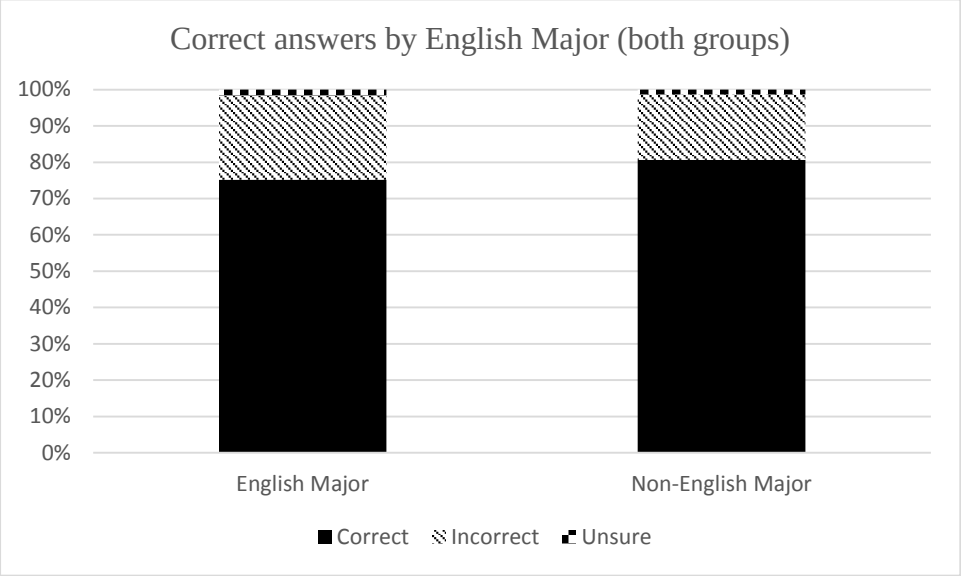

Table 7: Scores divided by English Major.

6.5 Comparison with Shimizu and Dantsuji's results

The following is a list of pairs presented in both this study and Shimizu and Dantsuji (1983)'s. The percentages represent the amount of respondents who answered a given word correctly. The values presented on the left side and the middle are my own, comprising of Group 1 and 2 respectively (see APPENDIX), with those on the right being Shimizu and Dantsuji's data. Correct answers have also been underlined.:

Right/Light 84.4%/64.3%/87.5%

Pirate/Pilot 81.3%/57.1%/59.4%

Read/Lead 93.8%/78.6%/84.4%

Raw/Law 75%/53.6%/84.4%

Pray/Play 96.9%/64.3%/75%

Rice/Lice 100%/82.1%/84.4%

Wrong/Long 81.3%/53.6%/90.6%

Mean: 87.5%/64.8/%80.8%

While the pronounced phoneme and the accompanying sound file(s) were not provided by Shimizu and Dantsuji, it is interesting to observe the considerable difference of some pairs among the three groups; Wrong/Long had a disparity of 37%, with Pray/Play not far behind at 32.6%. The former also only observed students of English literature, while this study observes non-English majors too, yet respondents in Group 1 still have more overall accuracy, but Group 2's is lower.

7. Discussion

7.1 Research Question #1

Does living in Sweden produce significantly different results?

Group 1 (Sweden) scored fairly well, with an overall accuracy of 86%.
Group 2 (Japan) scored notably lower than Group 1, with an average of 69.3%. The majority of respondents in said group showed an overall lower competency, pointing to long-term residents of Sweden as having higher perceptive ability. Despite this, it is difficult to pinpoint what the reason for these results is. Perhaps it is simply increased exposure to foreign language input, or even English itself. The two groups also had similar results in all other areas. An example of this is Figures 2 (a) and (b), in which easy and difficult words were largely similar. These words contained both /l/ and /r/ phonemes, meaning that one cannot therefore deduct which phoneme is easier. Similarly, both groups had the same difficulty pattern in terms of word type, with type 2 (medial, single) being the easiest, and type 3 (medial, consonant cluster) being the hardest. These facts point to Group 1 (Sweden) doing better quantitively, but not qualitatively. One can therefore argue that the differences between /l/ and /r/ phonemes in English vs Swedish therefore do not have much if any effect on a Japanese speaker's ability to perceive English. One can then conclude that living in Sweden long-term produces better overall ability, but does not enhance one's ability in any particular way.

7.2 Research Question #2

What is the role of phoneme position in a word?

What seems to be clear is that as predicted, Type 3 words are most difficult. While the Speech Learning Model (SLM) states that phonemes completely foreign to the learner are more likely to be easy to acquire for said learner, it does not speak for consonant clusters. One could apply the SLM to this and argue that since these (almost) do not exist in Japanese, they are easier to learn than their counterparts. The results, however, say the opposite. It is also not unheard of to hear Japanese speakers insert vowels into consonant clusters when speaking English, due to the lack thereof in their native language. When it comes to a given /r/ or /l/ phoneme's location in a word, said location being initial or medial does not appear to be significant, but rather what is, is whether it is surrounded by other consonants or if it is surrounded by vowels.

Also of note was the results of English majors vs non-English majors, in which the former consistently outperformed the latter to varying degrees (an 8% difference in Group 1, and a 3% difference in Group 2). While this may be a result of normal distribution, it could also point to English majors in general not having better ability to discern /l/ and /r/. Basic logic would lead one to believe that the English majors would have the upper hand when it comes to English, so the fact that they were not even equal but outright bested by their peers when one would believe the opposite is definitely of interest.

Though it can be said that Group 1 is overall better at audial perception, if the form were given to a native English speaker, it is doubtful that they would get many, if any answers wrong. A look at individual responses shows that some respondents had 100% accuracy, though these were very few. This points to problems existing even among those who have learned a language that most definitely contains [l].

7.3 Research Question #3

Is perception of /l/ greater than /r/ in group 1?

In short, no. There was a small difference favoring /r/, though this was likely a result of normal distribution. The overall better perception can be attributed to the longer and increased exposure to the /l/ and /r/ rather than any specific Swedish phonetics of /l/ and /r/ having improved a particular sound. There was not obvious influence from the Swedish (L2) phonetic and phonological patterns related to /l/ and /r/. Also, English is frequently used in Sweden, which speaks for 'more exposed' to the English environment.

7.4 Research Question #4

Do gender and English Major background affect perception?

The results provided by the data point to the answer being no, at least not to a decisive degree. Gender was of minor importance in Group 1, however in Group 2, females performed somewhat notably better overall. On the other hand, non-English majors fared better than their peers across the board.

7.5 Research Question #5

How can the results be used to further improve perception?

The lack of statistical significance (outside of the general disparity between the two groups) within the results makes it hard to draw conclusions in terms of solving this seemingly eternal problem. What is evident is that Group 1 outperformed 2, but the truly desirable result would be native-like perceptive ability. One could argue that mass immersion and simply "living the

language” would lead to proficiency, but there are many who do such and still cannot discern these sounds when presented with a minimal pair. On the other hand, there are individuals who do so to great effect. What the driving factor behind this is varies from person to person, but in an attempt to provide a means that is effective for the majority of individuals, one could claim that in order to solve this problem, phonetic training is required. A clear description coupled with images showing the locations of the tongue when pronouncing various phonemes can be of immense help, particularly with /l/ and /r/. Moreover, extra practice with words of particular difficulty (such as words with consonant clusters) could produce favourable results. While children are of course not shown tongue positions when learning their native language, they do not possess the latter until they receive enough input from adults around them. Adult language learners on the other hand, already have a native language and use said language as a base for acquiring almost any new information. It is also especially important to provide said education to younger children, who arguably have a much easier time acquiring new phonemes than adults. While complex terminology and such may not be of much use, simply having a native or otherwise very proficient speaker present coupled with pronunciation practice could work wonders.

7.6 Further research

Further research on this topic is definitely desirable. The actual perception of /l/ and /r/ has been tested a great deal on average Japanese speakers, and with this thesis, on non-English second language speakers too, so further analysis of perception for perception’s sake appears unnecessary, unless performed on such a large scale that it provides near-undeniable empirical evidence. Rather, what would seem beneficial would be to study how one can improve Japanese speakers’ perception and production of these particularly difficult

phonemes, and then perform tests to see the results. Effective methods could then perhaps be implemented into the English-teaching curriculum in Japanese schools, thereby reducing this problems' size considerably.

8. Conclusion

The perception patterns for the Japanese subjects in Sweden clearly performed better than those who are in Japan. The patterns shown by the two groups are very similar, both difficult words and easy words were nearly identical. Word medial position was easiest to identify while word initial consonant cluster was more difficult. The /l/ phoneme in word initial consonant cluster was the most difficult to identify. Neither phoneme is more or less inherently difficult to perceive than the other for the average Japanese speaker, instead they are a set; one that must be tackled with careful practice in both speaking and listening. The present study is a pilot study and is not without its shortcomings (no statistical treatment, gender disparity in Group 1, limited scope, number of respondents, etc), however, it does provide a set of empirical data that can be used for analysis, hopefully proving useful for any who wish to study this topic in the future.

9. References

Amanuma Y, Ōtsubo K, and Mizutani, O. (1989) *Nihongo Onsei-Gaku『日本語音声学』* [Japanese Phonetics]. Tokyo: Kuroshio Shuppan, pp 74-76.

Aoki, H. (1990). The pronunciation of Japanese by Americans (in Japanese)「アメリカ人による日本語の発音」, In *Nihongo to Nihongo Kyouiku『日本語と日本語教育』vol 3, Phonetics and Phonology* (in Japanese), pp 219-233. Meiji Shoin,

Aoyama, K et al. (2004). Perceived phonetic dissimilarity and L2 speech learning: the case of Japanese /r/ and English /l/ and /r/. In *Journal of Phonetics*, vol. 32, pp 233-250.

Bradlow, A et al. (1999). Training Japanese listeners to identify English /r/ and /l/: Long term retention of learning in perception and production. In *Perception & Psychophysics*, vol. 61 (5), pp 977-985.

Elert, C-C. (1966). *Allmän och Svensk Fonetik*. Stockholm: Almqvist & Wiksell, pp 40, 46-47, 52

Engstrand, O. (2004). *Fonetikens Grunder*. Lund: Studentlitteratur, pp 147, 152, 154-155, 167

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* (pp. 233–277). Timonium, MD: York Press

Goto, H (1971) Auditory Perception By Normal Japanese Adults Of The Sounds “L” And “R”. In *Neuropsychologia* vol. 9, pp 317-323.

Kitahara, Y and Uwano, Z. (2003). *Onsei・On’in『音声・音韻』*[Phonetics And Phonology]. Tokyo: Asakura Shoten, p 11.

Logan, J et al. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America* vol. 89 (2) pp 874-86

Maki-jita まき- じた【巻(き)舌】国語辞書 – *Daijisen『大辞泉』*https://dictionary.goo.ne.jp/jn/207368/meaning/m0u/ [accessed 2018-12-25]

McMahon, A. (2002). *An Introduction to English Phonology*. [ebook] [s.l.] Edinburgh University Press, pp 18-20 29, 32.

Miyawaki, K et al. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. In *Perception & Psychophysics*, vol. 18 (5) pp 331-340.

Nagano-Madsen, Y. (2005). Swedish learners (in Japanese)「スウェーデン語」in *New Dictionary of Japanese Education 『新日本語教育辞典』*, Nihongo Kyouiku Gakkai, Tokyo, Taishukan Shoten

Nagano-Madsen, Y., Okamoto-Gustafsson, Y., and Shimizu, Y. (2014). Acquistion of L2 Japanese by Swedish learners - The effect of a study stay in Japan on grammar and speech (in Japanese)「スウェーデン人学習者による日本語の習得―文法と音声の習得にみる日本留学の効果」. In *Journal of International Student Education* , vol 8, pp77-102, Kochi University

Ra-gyou「ラ行」Wikipedia https://ja.wikipedia.org/wiki/%E3%83%A1%E3%82%A4%E3%83%B3%E3%83%9A%E3%83%BC%E3%82%B8) [accessed 2018-12-25]

Shimizu, K and Dantsuji, M (1983). A Study on the Perception of /r/ and /l/ in Natural and Synthetic Speech Sounds. In *Studia phonologica*, vol. 17, pp 1-14.

Sugito, M (1984) *Physiological characteristics of za-, da-, and ra-gyou on, -Comaprison between Osaka and Tokyo speakers as well as American speakers.『ザ、ダ、ラ行音の生理的特徴 - 大阪、東京方言話者と アメリカ人の発音の比較による』*. Osaka Shoin

Womens college, NII-Electronic Library Service.
file:///C:/Users/xnagya/Desktop/Sugito.pdf

Riad, T. (2013). *The Phonology of Swedish*. [ebook] [s.l.] Oxford University Press, pp 57, 67-68.

Vance, T.J. (2008). *The Sounds of Japanese*. New York: Cambridge University Press, pp 17, 89.

10. Appendix

Table 1: Perception data for Group 1 cf. Figure 2(a)

Below is a list of the percentages of respondents who answered correctly on a given question, taken from the above chart. The mean value was 86.3% correct answers per question.

#1: Right/Light (S&D) (1) (L) 84.4%

#2: Pirate/Pilot (S&D) (2) (R) 81.3%

#3: Read/Lead (S&D) (1) (L) 93.8%

#4: Correct/Collect (2) (R) 100%

#5: Crime/Climb (3) (L) 65.6%

#6: Raw/Law (S&D) (1) (R) 75%

#7: Pray/Play (S&D) (3) (R) 96.9%

#8: Jerry/Jelly (2) (L) 90.6%

#9: Grass/Glass (3) (L) 90.6%

#10: Rice/Lice (S&D) (1) (R) 100%

#11: Arrive/Alive (2) (L) 84.4%

#12: Wrong/Long (S&D) (1) (R) 81.3%

#13: Berry/Belly (2) (L) 87.5%

#14: Bruise/Blues (3) (R) 75%

#15: Fry/Fly (3) (R) 87.5%

Table 2: Perception data for Group 2 cf. Figure 2(b)

#1: Right/Light (S&D) (1) (L) 64.3%

#2: Pirate/Pilot (S&D) (2) (R) 57.1%

#3: Read/Lead (S&D) (1) (L) 78.6%

#4: Correct/Collect (2) (R) 78.6%

#5: Crime/Climb (3) (L) 50%

#6: Raw/Law (S&D) (1) (R) 53.6%

#7: Pray/Play (S&D) (3) (R) 64.3%

#8: Jerry/Jelly (2) (L) 85.7%

#9: Grass/Glass (3) (L) 64.3%

#10: Rice/Lice (S&D) (1) (R) 82.1%

#11: Arrive/Alive (2) (L) 82.1%

#12: Wrong/Long (S&D) (1) (R) 53.6%

#13: Berry/Belly (2) (L) 78.6%

#14: Bruise/Blues (3) (R) 71.4%

#15: Fry/Fly (3) (R) 75%

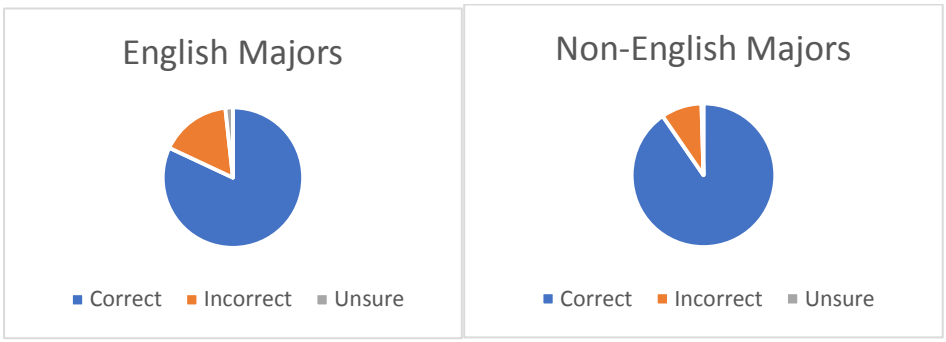

Figure 7: Background for Group 1 (Sweden)

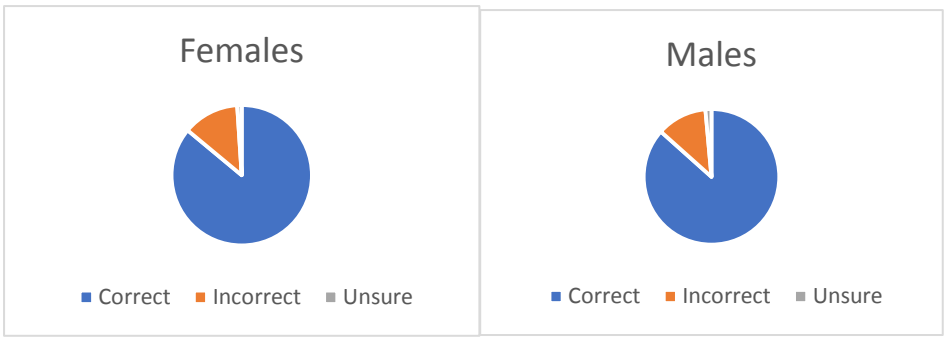

Figure 8: Gender difference for Group 1 (Sweden)

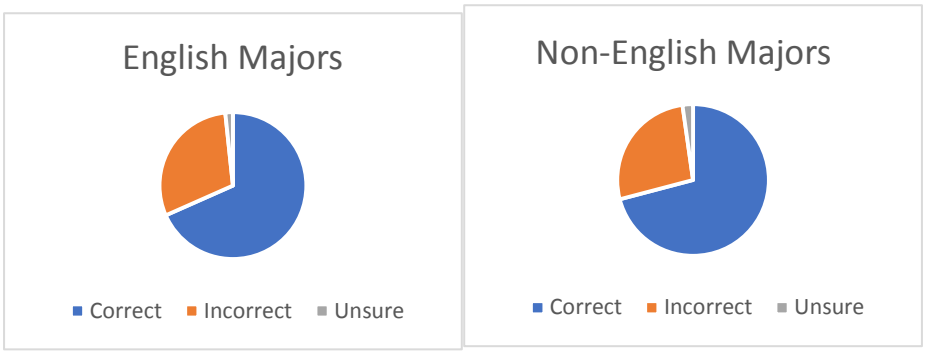

Figure 9: Background for Group 2 (Japan)

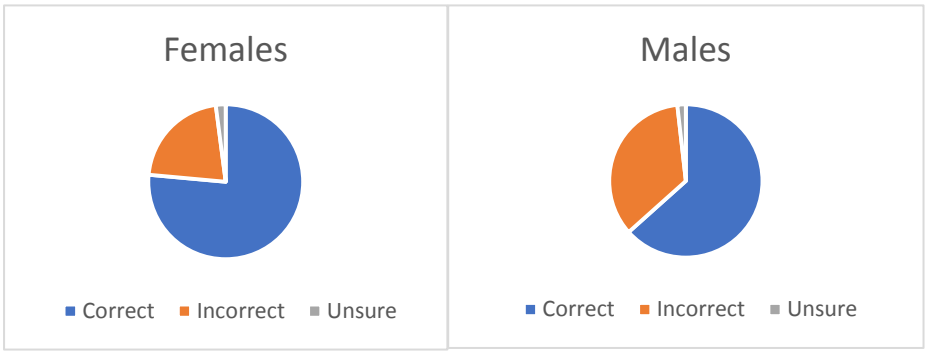

Figure 10: Gender difference for Group 2 (Japan)