# Examination of the [si] and [ʃi] Confusion by Japanese ESL Learners

by

Akitsugu Nogita

B.A., Aoyama-Gakuin University, 1999

A Thesis Submitted in Partial Fulfillment

of the Requirements for the Degree of

MASTER OF ARTS

in the Department of Linguistics

©Akitsugu Nogita, 2010
University of Victoria

All rights reserved. This thesis may not be reproduced in whole or in part, by photocopy or other means, without the permission of the author.

## Supervisory Committee

# Examination of the [si] and [ʃi] Confusion by Japanese ESL Learners

by

Akitsugu Nogita

B.A., Aoyama-Gakuin University,1999

## Supervisory Committee

Dr. Hua Lin, Supervisor
(Department of Linguistics)

Dr. Li-Shih Huang, Departmental Member
(Department of Linguistics)

**Supervisory Committee**
Dr. Hua Lin, Supervisor
(Department of Linguistics)

Dr. Li-Shih Huang, Departmental Member
(Department of Linguistics)

## Abstract

It is a general belief in Japan that the English /s/ and /ʃ/ before high front vowels (as in *see* and *she*) are problematic for Japanese ESL (English-as-a-second-language) learners. Some research has also reported the /s/ and /ʃ/ confusion by Japanese ESL learners. Their pronunciation errors are often explained based on phonetics, but there are reasons to believe that the learners' knowledge of the phonemes of the target words is at fault. This study examines 1) whether monolingual Japanese speakers distinguish the [si] and [ʃi] syllables in both perception and production in the Japanese contexts and 2) what would be the sources of Japanese speakers' challenges in mastering the distinction between [si] and [ʃi] in their English production if Japanese speakers can produce and perceive the difference between these syllables. This study conducted two experiments. In the first experiment, 93 monolingual Japanese speakers between the ages of 17 and 89 in and around Tôkyô read aloud the written stimuli that had [si] and [ʃi] in the Japanese contexts, repeated the sound stimuli that had [si] and [ʃi] in the Japanese contexts, and listened to the [si:] and [ʃi:] syllables in isolation recorded by a native speaker of Canadian English. The results showed that the participants all distinguished [si] and [ʃi] in both perception and production regardless of their ages. Based on these results, I hypothesized that the [s] and [ʃ] confusion by Japanese ESL learners is caused by misunderstanding, rather than an inability to articulate these sounds. In the second experiment, 27 Japanese ESL students

were recorded reading an English passage. The passage contains /s/ (7 times) and /ʃ/ (11 times) before high front vowels. After the reading, the participants were taught the basic English phonological system and the symbol-sound correspondence rules such as “s”-/s/ and “sh”-/ʃ/. The lesson lasted 40 minutes during which the participants were also interviewed to find out their awareness of the symbol-sound correspondence. No articulation explanations were given during the lesson. After the lesson, the participants read the same passage. The results showed that /s/ and /ʃ/ were mispronounced 39 and 67 times respectively in total by the 27 participants before the lesson, but only 7 and 19 times after the lesson. These changes are statistically significant. Moreover, the interview during the lesson revealed that the participants lacked phonological awareness in English as well as the knowledge of the symbol-sound correspondence rules. This study concluded that many of the mispronunciations by Japanese ESL learners, including /s/ and /ʃ/, can be solved by teaching the English phonics rules and some basic phonological rules without teaching the articulation of these sounds.

TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

# ACKNOWLEDGEMENTS

Without the knowledge, advice, support, and wisdom of great many people, the work presented here would not have been possible. First among these is Dr. Hua Lin, my thesis supervisor. Her support and guidance in my research and expertise in second language phonetics and phonology have been an invaluable source of motivation and inspiration. I also wish to thank Dr. Li-Shih Huang, my committee member, for guidance and enthusiasm about my research, and constantly offering her expertise in applied linguistics. A special thank you goes to Dr. Tim Iles, my external examiner, for his insightful feedback and suggestions. I thank Dr. Sonya Bird, Dr. Dave McKercher, and Dr. Ronald Wardhaugh for valuable linguistic information for my thesis. I thank Chris Coey for preparing the equipment for my experiments. I am also grateful to Dr. Tsutomu Akamatsu, Dr. Jun Arimoto, and Dr. Kazuo Nakayama for permitting me to use their figures and tables, and to Dr. Toni Borowsky for advice of phonology. Many thanks to all the UVic faculty and staff members whose support has been invaluable.

I wish to extend my sincere gratitude to UVic graduate students. I wish to thank Shu-min Huang for revising and editing my thesis, Arianne Truong and Nicholas Travers for judging the experimental data, Scott Moisik for recording the sound stimuli, Izabelle Grenon and Thomas Magnuson for all of their advice. I thank many others for their support. I would also like to thank my family members. I thank my wife, Hiroko, for recording the sound stimuli, helping me organize the experiments, and finding the participants. I thank my parents and my sister for all their help and support. Equally important to this study has been the cooperation by my about 150 participants who gave up their valuable time to participate in this project.

Finally, I would also like to acknowledge the institutions that provided me with

financial support over the course of my M.A. program. I thank the Department of Linguistics for employing me as a Teaching Assistant, the University of Victoria for an M.A. Fellowship (2008-2010), and Henry and Michiko Warkentyne for a Graduate Scholarship (2010).

# Chapter One

# INTRODUCTION

## 1.1 Study Motivation

Can native Japanese speakers *pronounce* the syllables [si] and [ʃi]? Can they *perceive* the difference between the syllables [si] and [ʃi][1]? These are among the main questions that are addressed in this thesis. The need to answer these questions comes first from the fact that many Japanese learners of English have trouble distinguishing the English /s/ and /ʃ/ before high front vowels.[2] For instance, many English pronunciation textbooks in Japan name the [s] and [ʃ] distinction as one of the major pronunciation problems of Japanese learners of English. Some research has also found this to be the case. For example, Fukushima, O'Connor, and Matsuzawa (2008) pointed out three major problems of Japanese speakers' English pronunciation as being the distinction between [s] and [ʃ], the distinction among [æ], [ʌ] and [ɑ], and the distinction between [ɹ] and [l]. The authors also placed the [s] and [ʃ] distinction at the same difficult level as the well-known [ɹ] and [l] confusion for the Japanese learners.

Why do Japanese learners of English confuse between English /s/ and /ʃ/? Common practice is to attribute the problem to the Japanese phonological system, where [s] and [ʃ] are said to be in *complementary distribution*: the phonological rule is that /s/ is realized as [ʃ] before /i/ and as [s] elsewhere (Masuda, 2003; Ohata, 2004). According to this phonological explanation, the [si] sequence does not exist in Japanese. If this is the case,

---

[1] When phonetically accurately transcribed, [ɕ], voiceless lamino-alveolo-palatal fricative, is often used in Japanese. Since this paper focuses more on phonology, I use [ʃ] in order to avoid complication. See 2.6 for the details.

[2] The defining characteristic of high front vowels is that the tongue is positioned high and front. For example, English has two high front vowels, /i/ as in "eat" and /ɪ/ as in "it."

Japanese ESL/EFL (English as a Second Language/English as a Foreign Language) learners are expected not to have difficulty in saying the word "she," but to have difficulty in saying the word "see." However, the problem seems much more complex than this. TuZino (2008), for example, observed that Japanese speakers often pronounce [si] in place of [ʃi]. This seems to suggest that Japanese ESL/EFL learners are expected not to have difficulty in saying the word "see," but to have difficulty in saying the word "she." A third group of researchers, such as Mabuchi (1971), Furuta, Suzuki, & Yamaguchi (1980), Yoshida (1982) and Quackenbush & Ôso (1990), believe both /si/ and /ʃi/ exist in the Japanese phonological system (as cited in Matsuzaki, 1993). While it seems to be generally said that older generations cannot pronounce some foreign originated sounds, including [si], the fact that nearly 40 years ago there was a scholar who phonologically separated /si/ from /ʃi/ is worthy of note. Likewise, TuZino phonologically separated [si] from [ʃi], although he grouped [si] and [sʷi][3] into /si/. In short, the status of [si] and [ʃi] in Japanese is under much debate. However, the question of whether both [si] and [ʃi] can be pronounced and distinguished by monolingual Japanese speakers has not been systematically and rigorously addressed. This present study aims to fill this gap by examining the status of /s/ and /ʃ/ in the Japanese language. One related question to answer in this thesis is whether Japanese ESL/EFL learners really need to practice the distinction between [s] and [ʃ] before high front vowels in both articulation and perception; more specifically, the thesis will investigate whether Japanese speakers really find the [si] and [ʃi] distinction problematic. If both [si] and [ʃi] exist in Japanese, Japanese ESL/EFL learners are not supposed to have difficulty in the English minimal pairs like "see" and "she," and the [s]

[3] [sʷ] indicates that [s] and [w] are produced quickly.

and [ʃ] confusion is not supposed to be caused by an inability to articulate, but by some other reasons.

## 1.2 Research Questions

The present study is designed to address the following four questions:

1. What phonological status does [si] have in Japanese?
2. Can monolingual Japanese speakers produce both [si] and [ʃi]?
3. Can monolingual Japanese speakers distinguish the [si] and [ʃi] syllables through listening?
4. If the Japanese speakers can produce and perceive the difference between [si] and [ʃi], what then would be the sources of Japanese speakers' challenges in mastering the distinction in their English production?

## 1.3 Organization of This Thesis

This thesis consists of five chapters. Chapter Two provides the answer to research question 1 by reviewing literatures regarding the background of Japanese phonetics and phonology, previous studies on Japanese ESL/EFL learners' /s/ and /ʃ/ confusion, and characteristics of English pronunciation textbooks in Japan. Chapter Three provides the answers to research questions 2 and 3 by reporting on an experiment on how monolingual Japanese speakers distinguish [si] and [ʃi] in both perception and production. Chapter Four provides the answer to research question 4 by describing an experiment on whether English phonological and orthographical knowledge alone helps Japanese ESL learners correct their errors regarding [s] and [ʃ] in the English context. Chapter Five concludes this thesis by providing a summary of the present study, pedagogical implications, limitations, future research directions, and the answers to all the research questions.

# Chapter Two

# LITERATURE REVIEW

## 2.1 Distribution of [s] and [ʃ] in Japanese

When Japanese phonology is addressed, different classes of lexicon often have to be considered separately, such as native Japanese words, Sino Japanese words (Chinese originated words usually written with *Kanji*), and foreign originated words usually written in *Katakana.*[4] According to the explanation of the model by Itô and Mester (1995), core lexical items satisfy all constraints; the less central areas the lexical items move towards, the more constraints they violate. In native and established loans (Chinese originated words) in modern standard Japanese, [s] usually does not appear before [i], and [ʃ] usually does not appear before [e]. Besides, even in loanwords, [ʃ] before [e] used to be depalatalized (Itô & Mester, 2003). For example, "**sh**epherd" was adapted to セパード [**s**epa:do]. (although these days both セパード [**s**epa:do] and シェパード [**ʃ**epa:do] are acceptable.[5]) On the other hand, [s] and [ʃ] have long been contrastive before [a], [ɯ], and [o] (Takayama, 2003): [kasa] 傘 "umbrella" versus [kaʃa] 貨車 "freight train," [kasɯ] 粕 "sediment" versus [kaʃɯ] 歌手 "singer," and [kaso] 過疎 "depopulation" versus [kaʃo] 箇所 "spot." In recent loanwords, in contrast, [ʃ] appears before [e] (Itô & Mester, 2003). For example, [se:kɯ] 成句 "idiom" and [ʃe:kɯ] シェーク "shake" are distinctive.

---

[4] In Japanese, there are two types of syllabaries (*Hiragana* and *Katakana*) and the logographic Chinese characters (*Kanji*).

[5] In this paper, phonological interpretation and transcription of Japanese words are based on Shinmeikai Japanese Accent Dictionary (Akinaga & Kindaichi, 2008). The long vowel symbol "—" in this dictionary is transcribed as /:/, the moraic placeless nasal is transcribed as /N/. I also sometimes use ↓ and ↑ as phonological pitch patterns based on this dictionary, but I omit them when I do not need to talk about pitch. This dictionary also shows vowel devoicing and the difference between [g] and [ŋ] but I omit them in my phonemic transcription since those are not phonologically distinctive.

To refer to orthography, according to Agency of Cultural Affairs, on June 28 in 1991 the Ministry accepted 13 new syllables in the Japanese writing system, as shown in Figure 2.1.

| シェ | チェ | ツァ | ツェ | ツォ | ティ | ファ | フィ | フェ | フォ | ジェ | ディ | デュ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ʃe | tʃe | tsa | tse | tso | ti | ɸa | ɸi | ɸe | ɸo | dʒe | di | djɯ |

***Figure 2.1.*** **13 new Japanese syllables accepted as official syllables in the writing system (June 28, 1991)**

In the writing system, "シェ," which represents [ʃe], is now officially accepted. Conversely, Agency of Cultural Affairs explained that it did not institute any rules for writing other special sounds, which include スィ [si], and some other syllables, such as ズィ [dzi], ニェ [nje], フョ [ɸjo]. This is a rule of the writing system, but it must be strongly related to the phonological system for it to be meaningful for our analysis. Table 2.1 shows the phonotactics[6] regarding /s/ and /ʃ/ and the *Katakana* symbol(s) for each syllable.

**Table 2.1**
***Japanese syllabaries regarding the consonants /s/ and /ʃ/***

| | **a** | **i** | **ɯ** | **e** | **o** |
|---|---|---|---|---|---|
| | サ | (スィ) | ス | セ | ソ |
| **s** | /sa/ | (/si/[7]) | /sɯ/ | /se/ | /so/ |
| | シャ | シ | シュ | シェ | ショ |
| **ʃ** | /ʃa/ | /ʃi/ | /ʃɯ/ | /ʃe/ | /ʃo/ |

## 2.2 *Chokuon* and *Yôon*

This section presents the difference between *Chokuon* [tʃo↑kɯ↓oN] and *Yôon* [jo↓ːoN] because these are related to lexical classes. There are a few variants of definition for these

[6] Phonotactics deals with restrictions in a particular language on the permissible combinations of phonemes.
[7] In this paper, I sometimes use the phonemic bracket / / for /si/ and /ʃi/ in Japanese to represent underlying representations, but it is not clear to what extent /si/ and /ʃi/ have a role in changing lexical meanings in Japanese.

terms.[8] From a perspective of the writing system, *Chokuon* refers to a syllable written with one symbol and *Yôon* refers to a syllable written with two symbols.[9] In this paper, I use this definition. The distinction between *Yôon* and *Chokuon* is considered to have originated from the influx of an overwhelming number of Chinese loanwords (Takayama, 2003). *Yôons* are used to spell the sounds that were originally not in the Japanese sound system. When Japanese needed new ways to write foreign originated sounds, Japanese speakers economically combined an existing big symbol and a small symbol to represent the new syllables (e.g., シャ [ʃa]) instead of creating new symbols (Inozula & Inozuka, 2009). In this paper, symbols for [sa], [ʃi], [sɯ], [se], and [so] are *Chokuon*, whereas those for [ʃa], [si], [ʃɯ], [ʃe], and [ʃo] are *Yôon*, as given in Table 2.2.

**Table 2.2**
***Classification of Chokuon and Yôon from the orthographic point of view***

| *Chokuon* | サ | シ | ス | セ | ソ |
|---|---|---|---|---|---|
| | /sa/ | /ʃi/ | /sɯ/ | /se/ | /so/ |
| *Yôon* | シャ | (スィ) | シュ | シェ | ショ |
| | /ʃa/ | (/si/) | /ʃɯ/ | /ʃe/ | /ʃo/ |

In Table 2.2, *Chokuon* (1 symbol – 1 syllable) is a syllable of the combination of /s/ and a vowel, and *Yôon* (2 symbols – 1 syllable) is a syllable of the combination of /ʃ/ and a vowel, except the combination with /i/ is in the opposite way. There are a few such complexities in the relationship between the writing system and the sound system in Japanese. The reason that the *Chokuons* サシスセソ [sa, ʃi, sɯ, se, so] consist of two different onsets[10]

---

[8] From a phonological point of view, *Chokuon* and *Yôon* are often considered as "plain" and "palatal" respectively (Itô & Mester, 2003). According to this definition, syllables like スィ [si] and ツァ [tsa] are not considered as *Yôon* because they do not involve palatalization.
[9] For example, in dictionaries such as "大辞泉 増補：新装版" and "数え方の辞典," syllables like スィ [si] and ツァ [tsa] are treated as *Yôons*. In this paper, I follow this definition.
[10] The Japanese symbols are organized in the 5×10 grid, 50 音表 "50-Sound-Chart." In this

seems to be explained in several different ways. According to Komatsu (1981), it is not the case that the シ /si/ is palatalized to [ʃi]. Instead, サシスセソ used to be pronounced as [ʃa, ʃi, ʃɯ, ʃe, ʃo]. In fact, based on J. Rodriguez Tçuzu's observation back in 1604, /se/ used to be pronounced like [ʃe], and the depalatalized [se] was Kantô accented pronunciation (Takayama, 2003). Eventually, [ʃ] became [s], but the シ [ʃi] retained [ʃ] because of the high front vowel [i]. This caused the inconsistency of the consonant quality in サシスセソ [sa, ʃi, sɯ, se, so]. Ôno (1994) mentioned that, recently, young girls' [ʃi] has been changing phonetically closer to [si], which illustrates the trace of the shift from [ʃ] to [s]. It is also said that the consonant of サシスセソ, namely /s/, used to be an affricate further back in history (Takayama, 2003).

### 2.3 The Status of [si] in Japanese

Does /si/ exist in Japanese? There are different views regarding this question. The [si] syllable is used in both written and spoken language in Japanese.

In the written language, the symbols "スィ," which are usually pronounced as [si], can often be seen recently. Although one might argue that this can be the evidence that the Japanese phonological system has /si/, the existence of certain writing ways does not necessarily bear direct relation to the existence or usage of /si/ and writing does not equal language (sounds or grammar). For example, the Japanese writing system sometimes uses the symbol "ヴ" to spell loanwords when the original words have [v]. However, it does not mean that Japanese owns the phoneme /v/, and, in fact, linguists usually do not accept /v/ in the Japanese phonological inventory for the reason that "ヴ" is mostly pronounced as [b] (e.g., Vance, 2008; Inozula et al., 2009). Another good example is that, as known as the

---

chart, the *Chokuons* サシスセソ [sa, ʃi, sɯ, se, so] are aligned in the same column.

issue of "*Yotsugana*" 四つ仮名, in the phonological system /dʒ/ and /ʒ/ are merged in most modern dialects, and so are /dz/ and /z/ (Takayama, 2003), but the writing system still retains ヂ versus ジ and ヅ versus ズ. However, these distinctions in the writing system do not represent the differences of the sounds, but more etymological aspects, as Japanese children are taught at school. To make matters more complicated, there are variants in symbols to represent loanwords. For example, when I typed "スィ" into CiNii Scholarly and Academic Information Navigator in October 2009, it gave me 570 hits, in which words that have "スィ" were mostly technical terms or foreign names. However, words like [**swi**ttʃi] スィッチ "switch" and [**swi**:tok:N] スィートコーン "sweet corn" have the symbols "スィ" which actually make the sound [swi], not [si]. Furthermore, TuZino (2008) mentioned that, in the modern writing system, "スィ" can be pronounced as either [si] or [sʷi], although [si] or [sʷi] are not phonologically contrastive in modern Japanese. Incidentally, TuZino added that in the historical *kana* orthography, [sʷi] can be written as "スヰ." It has long been discussed by many scholars that one (set of) symbol(s) can correspond to several different sounds while one sound can correspond to several different symbols in loanwords in Japanese. In other words, the symbols "スィ" might not always be pronounced as [si]; likewise, the syllable [si] might not always be written as "スィ." On the other hand, there must be a reason why the writing system has a way to distinguish between スィ [si] and シ [ʃi], as the aforementioned ヴ [v] and *Yotsugana* have the reasons of the distinctions in the writing system. It is worth examining why there is a way to distinguish between スィ [si] and シ [ʃi] in the writing system.

In the Japanese pronunciation seminar conducted by NHK[11] Communications Training Institute in July 2009, one instructor said that announcers are supposed to follow the established prescriptive standard, while they sometimes faithfully pronounce words based on the written language. Therefore, if the written language has "スィ," announcers would pronounce [si]. For example, one chief announcer at NHK pronounces [si] in the tennis word [passiNgɯ ʃotto] パッスィングショット "passing shot" when commentating on a tennis match. Therefore, the symbols スィ [si] are used to be differentiated from the symbol シ [ʃi] in sounds. However, in terms of phonological distinction, since [passiNgɯ ʃotto] パッスィングショット can also be replaced with [paʃʃiNgɯ ʃotto] パッシングショット without changing the lexical meaning in both spoken language and written language, this word cannot be evidence that [si] and [ʃi] are contrastive.

As for the spoken language, in the NHK Japanese pronunciation seminar in July 2009, one of the instructors mentioned that [si] may not be in the prescriptive standard language but different scholars define it differently, while another instructor mentioned that one could say Japanese has [si]. According to NHK (2005), these days, some syllables used only in foreign originated words, such as [tɯ], [dɯ], [si], [tsi] and so on, started to have status we cannot ignore. According to Matuzaki (1993), four (Mabuchi, 1971; Furuta, et al., 1980; Yoshida, 1982; Quackenbush, et al., 1990) out of 15 studies suggested that [si] is in the Japanese phonological system. In more recent interpretations, Inozuka and Inozuka (2009) treated /si/ as a Japanese syllable used in foreign words, while Vance (2008)

---

[11] NHK is Japan's national public broadcasting organization. The official English name is Japan Broadcasting Corporation.

separated /si/ from the syllables that are unanimously accepted.[12]

An interesting piece of research is Ishino's questionnaire back in 1981 cited by Matsuzaki (1992). Ishino asked both university students and members of society the following question: "When announcers pronounce the alphabet 'C' on TV or radio, which sound do you think is more appropriate, [si:] or [ʃi:]?" Ishino found that there were slightly more members of society who preferred [ʃi:] than those who preferred [si:], while there were many more university students who preferred [si:] than those who preferred [ʃi:]. Ishino said that at this rate, there was a high possibility that [si] was being established in the Japanese sound system.

Another interesting story is the comment from Inoue, Maruya, Ôno, and Ôoka (1990). Young Japanese girls tend to pronounce /ʃi/ as [si] in Japanese. Note that this is a different story from the observation that young girls' [ʃi] are phonetically becoming close to [si] as mentioned in 2.2. When Inoue et al. told female students not to pronounce /ʃi/ as [si], they answered that if they pronounced /ʃi/ as [ʃi], their English pronunciation would be skewed. This implies that [si] is considered to be "fancy" and some Japanese speakers purposely replace [ʃi] with [si]. Again, the difference between [si] and [ʃi] changes only the style of the conversation.

These days, Japanese speakers sometimes substitute [sii] for /-ʃii/[13] in *i*-adjectives. For example, /oiʃii/ おいしい "tasty" is pronounced as [oisii]. One of the good examples is that a well-known Japanese singer used to pronounce /oiʃii/ おいしい "tasty" as [oisii] in his

---

[12] Incidentally, Vance (2008) accepted the affricate counterpart /tsi/ in the Japanese phonological system by showing words such as [eɾitsiN] エリツィン "Yeltsin."
[13] The Japanese *i*–adjectives used to be divided into two types, *ku*-type and *shiku*-type. *I*–adjectives finishing in *shii* /-ʃii/ used to be *shiku*-type (Irie, 2004).

cooking show. Another well-known Japanese singer also pronounced /oiʃii/ おいしい "tasty" as [oisii] in a TV commercial for instant noodles. Moreover, [oisii] is often spelled with *Hiragana* as "おいすぃ～," which is differentiated from "おいしい" [oiʃii]. Again, in this case, the replacement of [sii] for /-ʃii/ may imply certain connotations but does not change the lexical meaning.

To sum up, Japanese speakers seem to recognize [si] and [ʃi] as two different syllables, and their recognition can be seen in the written language as the symbols スィ [si] and シ [ʃi]. Specifically, replacing [ʃi] with [si] seems to be a trend. However, this replacement changes only the style of the conversation, not its lexical meaning.

## 2.4 Minimal Pairs between [si] and [ʃi]

The existence of minimal pairs is good evidence of distinctive phonemes. What makes matters complicated is that since in Japanese some syllables appear only in foreign-origin words, such as [di] and [ɸi], it is sometimes not easy to find minimal pairs (Matsuzaki, 1992). Moreover, some variations are observed in the same word. An example of a variation between [di] and [dʒi] is that the word that originated from "radical" can be pronounced or written as both [ɾadikaɾɯ] ラディカル and [ɾadʒikaɾɯ] ラジカル. Such examples can cause an argument that [di] and [dʒi] do not change the meaning of this word and these syllables are not phonologically distinctive in Japanese, but such a conclusion is too simplistic. For example, in words like [diɾekɯtaː] ディレクター (director), [di] cannot be replaced with [dʒi] and *[dʒiɾekɯtaː][14] is impossible. Except for some words, [di] consistently occurs. For example, [diNkɯsɯ] ディンクス "double income no kids" and

[14] The symbol * indicates "unacceptable".

[dʒiNkɯsɯ] ジンクス “jinx” are a minimal pair. In fact, all 15 studies in Matsuzaki’s (1993) paper indicated [di] and its voiceless counterpart [ti] are in the Japanese phonological system. Since consistency of a certain sound differs from word to word, researchers cannot claim that [di] and [dʒi] are allophonic or free variants by showing few examples; for instance, “radical” can be either [ɾadikaɾɯ] ラディカル or [ɾadʒikaɾɯ] ラジカル (Matsuzaki, 1993). In other words, [di] and [dʒi] in these words could be purely personal, dialectal, or stylistic.

As for [si] and [ʃi], Matsuzaki (1993) searched 16 major dictionaries, and he found [si] in only two words in Sansêdô-Japanese-Dictionary (1982), [si:] スィー “the alphabet ‘C’” and [si:dzɯN] スィーズン “season.” These words have the symbols “スィ” which are usually pronounced as [si]. However, the dictionary also has the alternatives, [ʃi:] シー “C” and [ʃi:dzɯN] シーズン “season” respectively. The other 15 dictionaries Matsuzaki searched did not have any word that had “スィ [si].” [si] and [ʃi] do not seem to be contrastive in core lexical items at least until 1993. However, since dictionaries often aim to provide a “standard” which may differ from how words are produced in reality, the data from Matsuzaki (1993) are not sufficient for us to draw a conclusion (H. Lin, personal communication, 2010).

In peripheral lexical items, the issue of whether [si] and [ʃi] are distinctive is different. When Japanese speakers sing, they generally pronounce “do, re, mi, fa, so, la, ti, do” as [do, ɾe, mi, ɸa, so, ɾa, ʃi, do] ド, レ, ミ, ファ, ソ, ラ, シ, ド. Note that Japanese speakers do not pronounce “ti” as [ti] or [tʃi], but [ʃi] just like the Chinese do. However, Japanese speakers who specialize in vocal music pronounce “ti” as [si], which cannot be replaced with [ʃi]. For example, when students in music school in Japan take a vocal music

course, they are told not to pronounce “ti” as [ʃi]. More importantly, they still pronounce the Japanese word 死 “death” as [ʃi]. Therefore, in this cultural group, [si↓][15] “ti” and [ʃi↓] “death” behave like a minimal pair.

Another example is [heɾɯsi:] へるスィー. This is the name of a character in a comic book story called “*Takeshi*” written by Mitsutoshi Shimabukuro from 1997 to 2002. This name derives from the English word “healthy.” As far as I know, it is pronounced with [si:], which cannot be replaced with [ʃi:]. Moreover, Japanese has the loanword [heɾɯʃi:] ヘルシー “healthy.” Therefore, [heɾɯsi:] へるスィー and [heɾɯʃi:] ヘルシー could be a minimal pair. Although both words are from the same origin “healthy,” since the same original word can become lexically and phonologically different words, like “stick” becoming [sɯtikkɯ] スティック (a long thing) and [sɯtekki] ステッキ (cane) (Matsuzaki, 1992), that the origin is the same does not explain that [si] and [ʃi] are allophonic.

[baɾasi:] ばらスィー, the pseudonym of a Japanese comic author, provides another example. Here, [ɾaʃii] らしい “seem” and [baɾasi:][16] ばらスィー are a near minimal pair. Incidentally, what is intriguing is that [baɾasi:] ばらスィー is spelled with *Rômaji* (Japanese Romanization) as “*Barasui,*” in which [si:] is spelled as “sui.” This reason is not known, but this spelling may suggest the “glide-insertion” (see sections 3.2.6 and 3.3.7).

Another interesting word is [midzɯha**ʃi** paɾɯ**si:**] 水橋パル**スィ**, which has both [ʃi]

---

[15] Again, arrows indicate phonemic pitch patterns.
[16] I transcribe “ii” in [ɾaʃii] らしい as the vowel sequence [ii], instead of the long vowel [i:] because there is a morpheme boundary. There is a well-known minimal pair of a vowel sequence and a long vowel: /satooja/ 里親 “foster parent” and /sato:ja/ 砂糖屋 “sugar dealer” (Vance, 2008).

and [si], although [i] in [ʃi] is likely to be devoiced.[17] It is the name of a character of a computer game called "Tôhô-Project" in 2008. This game is popular among some young people, and it can be considered as a cultural dialect. When I informally asked some salespeople in several computer game shops how they pronounce 水橋パルスィ, all of them pronounced スィ as [si:], but not [ʃi:].[18] As well, when I informally asked those who like this game how they pronounce 水橋パルスィ on the website, I received eight replies. Two of them answered that they pronounced both [paɾɯʃi:] and [paɾɯsi:] although they knew that [paɾɯsi:] was proper, and the others answered either [paɾɯsi:] with a long vowel or [paɾɯsi] with a short vowel. According to this informal survey, using [ʃi] instead of [si] does not seem preferable. Moreover, replacing [ʃi] with [si] would never occur in this word: *[midzɯhasi paɾɯsi:] *水はすぃパルス. Intuitively, *[midzɯhasi paɾɯʃi:] *水はすぃパルシー sounds even stranger. This indicates that [si] and [ʃi] are not freely replaceable. All in all, in this lexical class, when the original sound is [si], it is preferable to keep it, rather than to palatalize it to [ʃi]. In other words, speakers are likely to be faithful to the original [si]. To describe how this word is pronounced, I use the tableau of the theoretical instrument from Optimality Theory, which can be illustrated as follows: the candidate (a) is optimal and the candidate (b) could be possible for some speakers but not preferable.

---

[17] In standard Japanese, [i] and [ɯ] between voiceless consonants are regularly devoiced (Vance, 2008).

[18] Whether "スィ" is [si:] with a long vowel or [si] with a short vowel is another complicated issue in Japanese (Miyauchi, 1993), which is beyond the scope of the paper.

| | /midzɯhaʃi paɾɯsi:/ | Ident I-O | Palatalization |
|---|---|---|---|
| a. | ☞ midzɯhaʃi paɾɯsi: | | * |
| b. | ? midzɯhaʃi paɾɯʃi: | *! | |
| c. | midzɯhasi paɾɯsi: | *! | ** |
| d. | midzɯhasi paɾɯʃi: | **! | * |

***Figure 2.2.*** **Tableau for /midzɯhaʃi paɾɯsi:/**
*Note*. Ident I-O = All the segments must be identical in both Input and Output; Palatalization = [s] must be palatalized before high front vowels; * = the candidate violates the constraint (When there are two symbols of *, the constraint is violated twice.); ! = the candidate is not optimal; ☞ = the candidate is optimal; ? = the candidate is not optimal but could be possible for some speakers. Ident I-O (faithfulness) outranks Palatalization (markedness).

To sum up, in central areas of the lexicon, no minimal pairs of /si/ and /ʃi/ are found. However, in peripheral lexical items, such as some cultural dialects and names, [si] behaves like a distinctive syllable, meaning that /si/ is a contrastive syllable of Japanese.

## 2.5 The Distinctive Function between [s] and [ʃ] in English

As stated by Rysiewicz (1990), in English, the alveolar segments /t, d, s, z/ become /tʃ, dʒ, ʃ, ʒ/ when followed by /j/; for example, the "s" in "expression" or "sensual" is pronounced as /ʃ/. This rule can apply across word boundary; for example, "s" in "miss you" can be pronounced as /ʃ/. Chomsky and Halle (1968) formulated that this rule applies when /j/ is followed by an unstressed vowel (as cited in Rysiewicz, 1990). Although this is the case most of the time, in fast or casual speech across word boundary, palatalization can occur when followed by a stressed vowel; for example, "s" in "this year" can be realized as /ʃ/. Moreover, /s/ and /ʃ/ can be neutralized when preceding unstressed /i/; for example, "c" in "associate," in which the "ci" syllable is unstressed, can be pronounced as either /ʃ/ or /s/ (T. Borowsky, personal communication, 2010; H. Lin, personal communication, 2010). Much like in Japanese, the discriminative function between /s/ and /ʃ/ in English becomes weaker

when before a palatal segment. In terms of L2 acquisition, for Japanese ESL learners, acquiring the English palatalization rule may be easy because of the cross-linguistic similarity, or it may be difficult because L2 learners may doubt the cross-linguistic similarity is real. Of course, in English, there are minimal pairs in core lexical items, such as "see" versus "she," so the discriminative function between /s/ and /ʃ/ before a palatal segment is stronger in English than in Japanese.

### 2.6 Phonetic Differences between English and Japanese

As I mentioned in footnote 1, the Japanese [ʃ] is better transcribed as [ɕ], voiceless lamino-alveolo-palatal fricative (Nakayama & Yamaguchi, 2003). According to Pan, Utsugi, and Yamazaki (2004), while the Chinese sound transcribed in Pinyin as "x" is always transcribed as [ɕ] in IPA, the Japanese counterpart has IPA variants between [ʃ] and [ɕ]. One of the possible reasons is that the Chinese "x" is articulated more back than the Japanese counterpart (Pan et al., 2004). Kazama, Machida, Matsumura, and Ueno (1993) have pointed out another possible reason that phoneticians who tried to transcribe the Japanese language with IPA considered the Japanese [ɕ] to be similar to the English [ʃ] and started to transcribe it as [ʃ], and it became habitual (as cited in Pan et al., 2004). Therefore, although [ɕ] is preferable, [ʃ] seems to be acceptable still. In this paper, I use [ʃ] for both Japanese and English in order to avoid complexity, except in cases where I need to describe phonetic differences.

Compared to [ʃ] in English, [ʃ] in Japanese is articulated further back and the front part of the tongue is higher (Pan et al, 2004). As well, [ʃ] in Japanese does not have lip rounding and dorsum elevation, unlike [ʃ] in English (Pan et al, 2004). Nakayama and Yamaguchi (2003) summarized the differences among [s], [ʃ], and [ɕ] in the point of

articulation; the point of articulation of [s] is "blade and alveolar ridge," that of [ʃ] is "blade and post-alveoli," and that of [ɕ] is "blade and palato-alveoli."

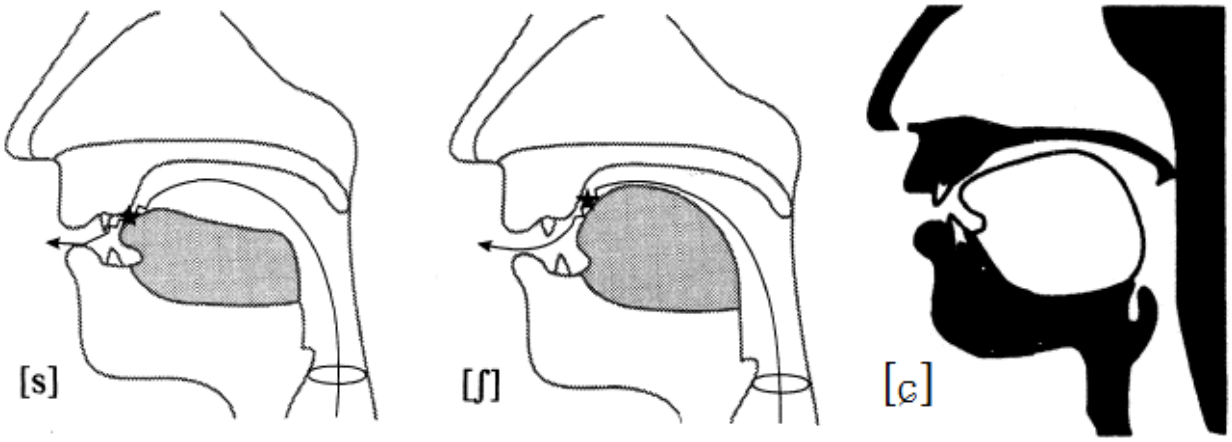


***Figure 2.3.*** **Comparison of articulation in [s], [ʃ] and [ɕ] (Nakayama & Yamaguchi, 2003, pp.214-216, reprinted with permission)**

Pan et al. (2004) analyzed the Japanese [ʃi], [ʃo], and [ʃa] and found that the energy of [ʃ] before [i] and [a] concentrates in 3000Hz to 6000Hz, or even higher by some individuals, and the energy before [o] concentrates in 3500Hz to 4000Hz. Arimoto (1993) showed the difference between the English "shoot [ʃu:t]" and Japanese "シュート[ʃɯ:to]" (shoot) in spectrogram as shown in Figure 2.4.

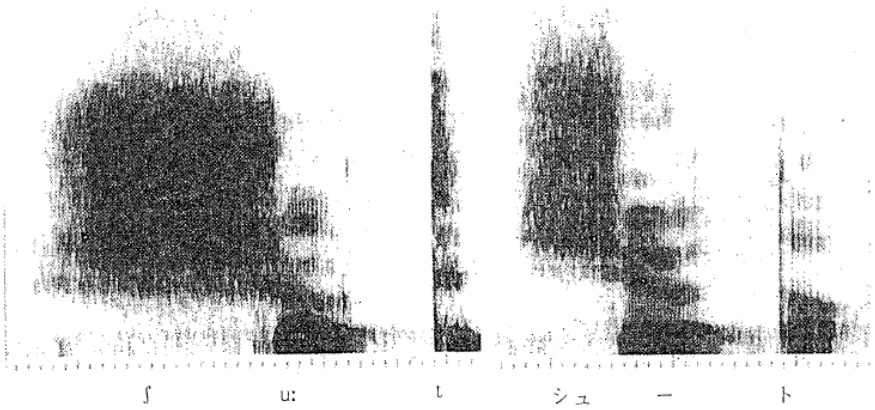


***Figure 2.4.*** **The English [ʃu:t] (left) versus the Japanese [ʃɯ:to] (right) (Arimoto, 1993, p.82, reprinted with permission)**

According to Figure 2.4, [ʃ] in Japanese is higher in frequency. Some may argue that [ʃi] pronounced by Japanese speakers are perceived as the English /si/ by English speakers. On the other hand, Arimoto (1993) said that teachers can decide whether they should teach the difference between English [ʃ] and Japanese [ʃ] to Japanese ESL/EFL learners since the Japanese [ʃ] is phonologically still within the English /ʃ/ (Eng/ʃ/) and it is not likely to impede communication.

Another phonetic characteristic of [ʃ] found by Pan et al. (2004) is that the frication of [ʃ] in Japanese is more steady than that in Korean, and the transition of the following vowel in Japanese is smaller than that in Chinese. This characteristic is notable in the [ʃa] sequence; therefore, we can interpret it as [ʃa] in Japanese, [ʃʰa] in Korean, and [ʃja] in Chinese with more detailed transcription.

As for the Japanese /s/ (Jap/s/), Arimoto (1993), who specializes in second language phonetics, said that although there were no solid data, a few students in a class seemed to pronounce /s/ like [θ] according to his impression. Arimoto further suggested that one of the major reasons is that the education of Japanese pronunciation to Japanese speakers had been neglected. Inozuka et al. (2009) have also mentioned that some young Japanese speakers tend to pronounce /s/ as [θ] but there is a lack of empirical evidence to support that statement. Inozuka et al. have assumed that when young Japanese speakers loosen their tongue tips, /s/ sounds closer to [θ]; therefore, [θ]-like /s/ is probably a type of weakened version of [s]. As indicated by Beckman, Edward, and Li (2009), Eng/s/ is clearly alveolar while Jap/s/ is more laminal and possibly somewhat dentalized. As well, Jap/s/ is less sibilant than that of English. For that reason, that Jap/s/ is likely to be toward [θ] is understandable. Beckman et al. (2009) also pointed out that children acquiring English

acquire /s/ earlier than children acquiring Japanese do. They suggested that adult English listeners who have no linguistic background may accept a wider range of centroid frequencies (centre of mass of the distribution in the fricative) for correct /s/ than adult Japanese listeners who have no linguistic background do. The way English speakers acquire /s/ may not be the same as the way Japanese speakers do. As for an acoustic characteristic of Jap/s/, Satô (2001) analyzed the Japanese [s] and [ʃ] in the #_[i] context and found that the energy of [s] concentrates around 4000Hz while that of [ʃ] spreads to the lower region (as cited in Pan et al. 2004).

Considering those phonetic differences, when Japanese ESL/EFL learners try to pronounce /ʃ/, it may be categorized as Eng/s/ by English speakers, and when Japanese ESL/EFL learners try to pronounce /s/, it may be categorized as Eng/θ/ by English speakers.

## 2.7 Previous Studies about Perception of English [si] and [ʃi] by Japanese Speakers

There are several pieces of research about how Japanese speakers perceive English [s] and [ʃ]. For example, Berman, Lambacher, Martens, and Nelson (2001) asked 104 Japanese first-year university students in Fukushima Prefecture to listen to 75 syllables presented in CV, VCV, and VC[19] syllables spoken within a varied-vowel environment by three phonetically-trained native English speakers. There were a total of 225 tokens (3 speakers × 75 stimuli), and the listeners identified each syllable as containing /f/, /s/, /ʃ/, /θ/, and /h/. The listeners had six years of prior English instruction at the junior and senior high school levels, but they had little or no exposure to training in English listening or pronunciation in a formal setting. A five-alternative, forced-choice answer containing

---

[19] C represents "consonant." V represents "vowel."

printed single letters in regular orthography (F, S, SH, TH, H) was passed out to each of the listeners. They found that /s/ and /ʃ/ were most confusable when preceded by /i/. As shown in Table 2.1, Japanese has the /s/ and /ʃ/ distinctions, except the [si] sequence is not as common as the others. Therefore, it is not surprising that the Japanese participants correctly heard the difference between /s/ and /ʃ/ with the exception of before high front vowels. Interestingly, Japanese speakers' ears are supposed to be used to the [ʃi] sequence, but when [ʃ] was presented in CV and VCV syllables combined with [i], more listeners incorrectly chose /s/ than those who correctly chose /ʃ/. The researchers analyzed that the increased confusability could have been caused by the listeners' bias toward /s/; when [ʃi] was presented, listeners may have disbelieved their own ears. It seems that psychological factors, such as over-generalization, are involved.

Another interesting study is the one conducted by Masuda (2003). The participants were 151 Japanese EFL learners who were second and third year university students majoring in English and took English phonetics classes in 2001 and 2002. They listened to 10 pairs of consonants, each of which was embedded in 10 sentences (100 sentences in total) containing minimal pairs. This was an ABX discrimination task[20]; a native English speaker pronounced each sentence three times, one of which was different from the other two, and the listeners chose the odd one. For example, if the speaker pronounced "Take a <u>seat,</u> please" the first and second time and "Take a <u>sheet,</u> please" the third time, the listeners were supposed to choose the third one. The results of /s/-/ʃ/ were unexpectedly good. In detail, in eight of the 10 stimuli, /s/ and /ʃ/ appeared before non-high-front vowels, such as

---

[20] In an ABX discrimination task, participants are presented with two known stimuli, A and B, and one unknown sample X.

"save" vs. "shave." Not surprisingly, they got correct answers over 81% of the time in those eight questions. This result agrees with the experiment by Berman, et al. (2001) where Japanese speakers have no difficulty in the /s/ and /ʃ/ distinction before non-high-front vowels. What is interesting is the other two pairs: "Take a seat, please," versus "Take a sheet, please," and "His seat is covered with dust," versus "His sheet is covered with dust." Unexpectedly, in the first pair, the listeners correctly answered 90% of the time. On the contrary, in the second pair, they got the correct answer only 59% of the time. Masuda analysed that as the result of blending, [z] in both "His seat" and "His sheet" disappeared and those sounded like [hɪsit], [hɪʃit], respectively. The listeners were so used to "spelling pronunciation" that they could not understand the meaning of the words when blending occurred, and they failed to guess it from the context. Masuda analysed that the participants' errors may be related to the participants' unfamiliarity with native speakers' reduced pronunciation. Therefore, his study implies that the general belief that Japanese ESL learners' /s/ and /ʃ/ confusion is due to the participants' inability to perceive the /s/-/ʃ/ distinction must be considered.

Adachi and Yamada's (1998) study involves Japanese university students and members of society. In their study, 22 consonants were embedded in both [C + ai] and [C + i] and 37 Japanese EFL listeners identified the consonants. The sounds [s] and [ʃ] were not confused in the [C + ai] context. This result agrees with the other two aforementioned experiments. On the other hand, the listeners were confused in the [C + i] context. On the basis of Adachi and Yamada's study, Nakayama and Yamaguchi (2003) proposed that native Japanese speakers are expected to transfer the palatalization of [s] before high front vowels to English and they may pronounce [si] and [sɪ] as [ʃi] and [ʃɪ]; for example, they are

expected to pronounce “see” as “she.” However, Nakayama and Yamaguchi also mentioned that they need to figure out why Japanese ESL/EFL learners also sometimes pronounce “ship” and “sheet” as “sip” and “seat” respectively; in other words, they are interested in finding out why [ʃ] is confused with [s]. In fact, according to Adachi and Yamada’s study, when the listeners heard [ʃi] they answered [si] nearly 30% of the time, which is a similar result to that of Berman, et al. (2001). If I assume that /s/ is palatalized to [ʃ] before /i/ in Japanese and Japanese ESL/EFL learners transfer it to English, this assumption cannot explain their confusion pattern.

These three studies all agree with the conclusion that native Japanese speakers misperceive Eng/s/ and Eng/ʃ/ before [i], but they correctly perceive the difference before other vowels. However, the details were not the same. In Berman et al.’s (2001) study, the participants incorrectly heard Eng/ʃ/ as /s/ more than 50% of the time, while in Adachi and Yamada’s (1998) study, the participants confused Eng/s/ and Eng/ʃ/ about 30% of the time, while in Masuda’s (2003) study, the participants correctly distinguished word-initial Eng/s/ and Eng/ʃ/ 90% of the time. Considering these results from the previous studies, it is necessary to reexamine the general belief that Japanese ESL learners cannot distinguish between [s] and [ʃ] before high front vowels because of the palatalization rule, /s/ → [ʃ] / _i, in Japanese.

## 2.8 Loanword Interference

Avery and Ehrlich (2003) mention that many of the mispronunciations by Portuguese ESL learners can be traced to the influence of the Portuguese spelling system rather than to an inability to produce particular sounds because Portuguese speaking teachers of English with a large number of Portuguese students would familiarize

themselves with the Portuguese symbol-sound correspondence rules. Might Japanese ESL learners' mispronunciation be a similar phenomenon? Goble (2002) observed Japanese college students in the classroom as both a teacher and as an observer of learning for years. As well, he has examined students' ability to distinguish English loanwords in Japanese and original English words. He also examined his students' ability to correctly pronounce and spell some English words that had been imported in the Japanese vocabulary. He explained, "The results of these tests showed an astonishing lack of awareness among students that Japanese loanwords and their English counterparts are not equivalent entities, while also revealing an inordinate amount of mother-tongue interference in students' L2 usage" (p.55). He also mentioned that Japanese students often find it difficult to accept that English words pronounced by native speakers are proper since the sounds of English loanwords in Japanese are already a long-standing part of their L1 linguistic knowledge. Loanword sounds seem deeply engraved in Japanese ESL/EFL learners' brains. In fact, my participants I will mention in Chapter Four were no exception. Goble also listed some of the students' spelling errors that were considered as loanword interference. Goble did not analyze the details of the errors because his focus was not on error analysis, so I picked up some interesting spelling errors from his paper and observed some details. In Table 2.3, I list the original English words, the equivalent loanwords in Japanese phonologically transcribed with IPA, and what the students actually spelled. The symbols "↑" and "↓" show the accent patterns in the Japanese words.

Table 2.3
*Japanese college students' misspelling (Goble, 2002, p.66-67)*

| Original English words | Japanese loanword pronunciation | Students' spelling |
|---|---|---|
| salad | /sa↓ɾada/ | sarada |
| doughnuts | /do↓:na(t)tsɯ/ | donatsu |
| volleyball | /ba↑ɾe:bo↓: ɾɯ/ | ballayball |
| tennis court | /te↑nisɯko↓:to/ | tennis coat |
| gallery | /gʲa↓ɾaɾi:/ | garaly |
| stadium | /sɯ↑ta↓dʒiamɯ/ | stagiam |
| concert hall | /ko↑Nsa:toho↓: ɾɯ/ | concert hole |
| platform | /pɯ↑ɾattoho↓:mɯ/ | platt home |
| necklace | /ne↓kkuɾesɯ/ | neckless |
| drum | /do↓ɾamɯ/ | doram |
| slipper | /sɯ↓ɾippa/ or /sɯ↑ɾi↓ppa/ | surippa |
| apron | /e↓pɯɾoN/ or /e↑pɯɾoN/ | eplon |

Some types of consonant confusion, such as "r" versus "l" and "b" versus "v," are not surprising since Japanese does not have these contrasts. What is interesting is that some contrasts which exist in Japanese or are similar to Japanese sounds were confused as well. For example, "d" and "g /dʒ/" in "stadium," "f" and "h" in "platform," "a /ei/" and "e /ɛ/" (see 4.12) in "apron" were mixed up. Loanword interference accounts for it. In the word "salad," although common epenthetic vowels after "d" to avoid a word-final obstruent are "u" or "o," "salad" adopted "a," which is a very rare epenthetic vowel in Japanese (Horiguchi, 1997). The influence of the loanword /sa↓ɾada/ サラダ "salad" also accounts for the students' error. Some consonant clusters, such as "st" in "stadium" and "pl" in "platform," were left not being broken apart by inserting a vowel, while "dr" in "drum" and "sl" in "slipper" were broken apart with a vowel: "*doram" and "*surippa." This is probably because the epenthetic vowels in these loanwords are accented and not devoiced; therefore, the epenthetic vowels are salient. Interestingly, those students' spelling choices are not the

same as the way of Japanese Romanization. For example, the word "stadium" is adapted to /sɯ↑ta↓dʒiamɯ/ スタジアム spelled as "*sutajiamu*" in Japanese, but the student spelled it as "*stagiam." In Japanese Romanization, the letter "g" always corresponds to /g/, so those students may have implicitly learned the English Soft-G. Goble (2002) considered these types of errors as "mis-application of . . . L2 spelling rules" or "inter-language transfer" (p.66). It is possible that pronunciation patterns by Japanese ESL/EFL learners is similar to these spelling error patterns.

## 2.9 Challenges of English Pronunciation Textbooks in Japan

The main motivation of this study is to examine whether the problem of the /s/ and /ʃ/ distinction before high front vowels might have been made up by misleading teaching. Bookstores in Japan carry a large number of English pronunciation textbooks. This indirectly reflects that many Japanese ESL/EFL learners have difficulties in English pronunciation or are interested in it. As well, each author's passion for reforming the English education in Japan can be detected from the books. On the other hand, Arimoto (2002) stated that some pronunciation textbooks were written based on the authors' experience of studying abroad and are not accurate, and other textbooks are misleading for beginners. Arimoto also mentioned that some practices in textbooks do not seem necessary. Makino (2008a) stated that quite a few authors of English pronunciation textbooks insist on the "wild method[21]" as a "new theory." I speculate that some of the Japanese ESL learners' weaknesses, including the /s/ and /ʃ/ distinction, might be induced by these teaching methods, as Richards (1971) called "false concepts" and Stenson (1974) termed "induced errors" (as cited in Brown, 2007). For example, as far as I have checked textbooks that

[21] In the original paper of Makino's (2009), he used the Japanese words *tondemo* トンデモ which means "unexpectedly crazy."

focused on individual English sounds, most of the textbooks point out the /s/ and /ʃ/ distinction as a problematic distinction and encourage learners to practice. One of the goals of this present study is to figure out whether the /s/ and /ʃ/ distinction really requires Japanese ESL/EFL learners to practice articulation.

In terms of second language education, I have read some questionable expressions in textbooks of English pronunciation for Japanese EFL learners, textbooks of the Japanese language, and even academic papers in journals of applied linguistics. Those expressions are not necessarily inaccurate, but they could give learners negative impressions on learning English pronunciation or may lead learners toward unnecessary articulatory practice. I will show some examples the misleading expressions.

**a.** English has far more consonants than Japanese (Ikegaya, 2008).

Since Ikegaya's (2008) textbook does not provide the consonant inventories, I refer to another paper. Takahashi (2005) showed both the 14-Japanese-consonant chart and the 24-English-consonant chart for the comparison of the cross-linguistic differences (p.52) as shown in Figure 2.5.

**English**

| | Bi-labial | Labio-dental | Den-tal | Al-veolar | Post-Al-veolar | Ve-lar | Glot-tal |
|---|---|---|---|---|---|---|---|
| Plosive | p<br>b | | | t<br>d | | k<br>g | |
| Affricate | | | | | tʃ<br>dʒ | | |
| Fricative | | f<br>v | θ<br>ð | s<br>z | ʃ<br>ʒ | | h |
| Nasal | m | | | n | | ŋ | |
| Lateral | | | | l | | | |
| Semi-vowel | w | | | r | j | (w) | |

**Japanese**

| | Bi-labial | Al-veolar | Post-Al-veolar | Ve-lar | Glot-tal |
|---|---|---|---|---|---|
| Plosive | p<br>b | t<br>d | | k<br>g | |
| Fricative | | s<br>z | | | h |
| Nasal | m | n | | | |
| Flap | | r | | | |
| Semi-vowel | w | | j | (w) | |

***Figure 2.5.*** **The inventories of English consonants and Japanese consonants (Takahashi, 2005)**

However, researchers do not all agree that Japanese has 14 consonants. For example, Tatsumi (2009) said that Japanese has 18 consonants. The Speech Accent Archive listed 20. Tsujimura (2007) listed 23. Akamatsu (2000) listed 27 as shown in Figure 2.6.

**Japanese Consonant Phonemes (Majority Speech)**

| | | | "labial" | "apical" | "palatal" | "hiss" | "hush" | "dorsal" | "liquid" |
|---|---|---|---|---|---|---|---|---|---|
| | | "voiceless" | /p'/ | | | | | | |
| | "palatalized" | | | | | | | | |
| | | "voiced" | /b'/ | | | | | | |
| "non-nasal" | | | | | | | | | |
| | | "voiceless" | /p/ | | | | | | |
| | "non-palatalized" | | | | | | | | |
| | | "voiced" | /b/ | | | | | | |
| | "palatalized" | | /m'/ | /n'/ | | | | | |
| "nasal" | | | | | | | | | |
| | "non-palatalized" | | /m/ | /n/ | | | | | |
| | | "voiceless" | | /t/ | | | | | |
| "non-nasal" | | | | | | | | | |
| | | "voiced" | | /d/ | | | | | |
| | | "voiceless" | | | | | | /X'/ | |
| | "palatalized" | | | | | | | | |
| | | "voiced" | | | | | | /Y'/ | |
| | | "voiceless" | | | | | | /X/ | |
| | "non-palatalized" | | | | | | | | |
| | | "voiced" | | | | | | /Y/ | |
| | "palatalized" | | | | | | | | /ʔ'/ |
| | "non-palatalized" | | | | | | | | /ʔ/ |
| | | "voiceless" | | | | /s/ | /ɕ/ | | |
| | | "voiced" | | | | /z/ | /ʑ/ | | |
| "fricative" | | | /ɸ/ | | /ç/ | | | | |
| "affricate" | | | | /ts/ | /cç/ | | | | |
| "spirant" | | | | | /j/ | | | /ɯ/ (/ɚ/) | |
| "glottal" /h/ | | | | | | | | | |

*Figure 2.6.* **The Japanese consonant inventory (Akamatsu, 2000, p.299, reprinted with permission)**

Akamatsu also provided another version of the inventory that has 29 consonants, in which the gradually disappearing /ŋ/ and /ŋʲ/ are included. Therefore, I could argue that 14, 18, 20, 23, 27, and 29 are all possible numbers of the Japanese consonant phonemes depending on the definition or the purpose. Interestingly, Tsujimura's 23 inventory includes /N/, while Akamatsu's 29 inventory does not include it because Akamatsu treated the moraic nasal as an archiphoneme, but not as a regular consonant. Therefore, the inventory with a bigger number may not necessarily cover all consonants listed in a smaller inventory. The point is that the 14-consonant inventory is not likely a truthful representation of the Japanese

phoneme inventory, which can be much larger depending on the researchers who study them.

The 14-consonant version provided by Takahashi (2005) has both advantages and disadvantages. The first advantage is that it looks simple, which satisfies one of the phonological modeling principles that the fewer phonemes, the better. The second advantage is that this inventory corresponds to the orthography-based 50-Sound-Chart.[22] In this inventory, for example, [ʃ] and [s] are grouped in /s/, [t], [ts] and [tʃ] are grouped in /t/, and [h], [ɸ] and [ç] are grouped in /h/. This grouping is useful in some morphological analyses since in some lexical classes, these sounds behave like allophonic variations. For example, the inflectional variations of the verb "lend," /kasa/ or /kaso/ (irrealis), /kasi/ (adverbial), /kasɯ/ (conclusive), and /kase/ (imperative), in which the stem is /kas/, are pronounced as [kasa], [kaso], [kaʃi], [kasɯ], and [kase] respectively. The disadvantage, on the other hand, is that this inventory is quite inconvenient when researchers transcribe some foreign originated sounds. For example, there are lexical contrasts between [tiː] ティー "tea" and [tʃii] 地位 "status", or between [ɸaiɾɯ] ファイル "file" and [haiɾɯ] 入る "enter." In the former pair, [t] and [tʃ] are irreplaceable, and so are [ɸ] and [h] in the latter pair. Even in native Japanese words, [gottsaN] "thank you (for a meal)," that contains [tsa], cannot be replaced with *[gottʃaN] or *[gottaN] (Kindaichi, 1989; Vance, 2008). If teachers and researchers assume that Japanese has 14 consonant phonemes and [ʃ] and [s] belong to /s/, [t], [ts] and [tʃ] belong to /t/, [h], [ɸ] and [ç] are grouped in /h/, and so on, there needs to be an extra explanation that Japanese speakers are aware of the phonetic differences of

[22] 50-Sound-Chart (50 音表) consists of five vowel letters on the vertical column and nine consonants on the horizontal row. Plus there are four *dakuon* 濁音 consonants and one *handakuon* 半濁音 consonant.

some allophonic variations, so some of the alternations, such as /t/ → [tʃ] / _i, do not occur in certain lexical classes, as Makino (2008a) provided. The second disadvantage is that in the comparison between the 24-English-consonant system and the 14-Japanese-consonant system, /ti/, for example, is pronounced as [ti] in the English contexts, but as [tʃi] in the Japanese contexts, which can be confusing. Due to these complexities, the 14-consonant inventory does not seem optimal among the various versions in terms of cross-linguistic comparisons. Moreover, since /ʃ/ is not included in the 14 consonants, there is no wonder that some people misunderstand that [s] and [ʃ] are not distinctive in Japanese.

**b.** Japanese does not have /f/, /v/, /θ/, /ð/, /ʃ/, /ʒ/, /tʃ/, or /dʒ/. English has far more consonants than Japanese (Ohata, 2004).

Again, a phonological analysis all depends on how researchers interpret a pair like [ti:] "tea" and [tʃii] "status." Although Ohata's analysis may be one of the possible ways, it does not seem optimal in terms of cross-linguistic comparison as mentioned above. Again, there is no wonder that some people misunderstand that in Japanese [s] and [ʃ] are not distinctive if seeing this expression in a paper of applied linguistics.

These examples are the tip of the iceberg. Ironically, many of the problems of the Japanese ESL learners might have been caused by those textbooks, but not by L1 constraints. In addition, Takahata (2003) mentioned that Japanese sport coaches tend to focus more on weaknesses than strengths. He suggested that coaches should encourage players to work on their strengths. Then, the coaches can tell the players that they can

improve their weaknesses as well because they have mastered their strengths this far. Likewise, English pronunciation materials in Japan tend to emphasize weaknesses. Those textbooks may need to focus more on the cross-linguistic similarities as Japanese learners' strengths. L. Huang (personal communication, 2009) gave an example of an encouraging way of teaching. Pronunciation lessons are often called "accent reduction," which means to reduce learners' non-native accent. Alternatively, the same lessons can be called "accent addition," which means to add a new accent to learners. The point is that the same issue can be seen from either a positive perspective or a negative perspective. To improve pronunciation textbooks in Japan, phonologists need to investigate a possibility that sounds that are generally believed to be weaknesses of Japanese ESL learners can be strengths if those are seen from different angles, and need to emphasize the strengths.

Besides, English pronunciation textbooks in Japan often provide details of articulation and acoustics. However, articulation is not the whole story of pronunciation problems (Fraser, in press). What is important for learners is to understand that phonemes change lexical meanings. To improve pronunciation textbooks in Japan, phonologists, especially those who specialize in second language phonology, may have to be involved more.

## 2.10 Perception of the Japanese Language

In English pronunciation textbooks in Japan, the maximally estimated number of English phonemes and the minimally estimated number of Japanese phonemes are often compared in order to conclude that Japanese pronunciation is far simpler than that of English (or any other language). In contrast, as far as I know, statements like "English pronunciation is far simpler than Japanese pronunciation" are not found. This tendency

may be derived from what Shiba (2008) called *an inferiority complex of Japanese*[23]; Japanese speakers tend to presume that the Japanese language is inferior to English (or any other language). In fact, the tendency of excessive underestimation of the Japanese language does not seem to stay within the pronunciation issues. For example, Kuwabara (2009), a professor at Kandagaigo University specializing in English syntax, stated in his presentation that Japanese speakers often say that Japanese is vague in terms of syntax while English is logical when it is linguistically not true. Another example can be found in Inoue, Marutani, Ôno and Ôoka (1989) who mentioned that students at school tend to see teachers of English "more fancy" and "more intelligent" than teachers of Japanese, and students think that Japanese is inferior to English. Inoue, et al. also pointed out the problem that teachers of Japanese neglect to teach the Japanese pronunciation and grammar to Japanese speakers at school. Another study (Shiba, 2008) mentioned that these days, even Japanese companies excessively use foreign originated words, instead of native Japanese words. Moreover, according to Shiba, in the past, Japanese people who intended to be linguists chose other languages as their specialties than Japanese because of the perceived inferiority of Japanese, which delayed the linguistic studies of the Japanese language by researchers. Shiba analysed that it probably resulted from the event that Japanese speakers imported Chinese characters when Japanese did not have its own writing system. The belief that Japanese is inferior to English (or any other language) may be deeply rooted in history. The followings are some examples of the inferiority complex of Japanese.

---

[23] Shiba (2008) uses the Japanese term "日本語コンプレックス" (p.12). I translated it into "an inferiority complex of Japanese."

a. The English sound system is far more complex than that of Japanese. Hence, native speakers of Japanese have a huge disadvantage of acquiring the English pronunciation (Fukasawa, 2007).
b. The Japanese language is very simple in terms of pronunciation (Fukushima, 2009).
c. The Japanese language has exceptionally few sounds among the languages around the world. Therefore, English has so many sounds that do not exist in Japanese (Haraga, 2009).

These expressions are all from English pronunciation textbooks and an article in a journal of the English education. From a linguistic point of view, it is not preferred to say that a certain language is easier than another because there is no single scale from easy to difficult, and in fact, pronunciation in most languages are considered equally difficult (Andersson, 1998). If one claims that Japanese pronunciation is an extreme exception, he or she would have to explain what motivation makes it exceptionally simple by providing evidence. Moreover, even if English has more or fewer phonemes than Japanese, the number of phonemes and difficulties of pronunciation may not be correlated (D. McKercher, personal communication, 2009; H. Lin, personal communication, 2010). Furthermore, the number of phonemes is just one of the many facets of pronunciation that include segmental factors such as sound inventory, phonotactic constraints, syllable structure, and so forth, not to mention many factors at the higher prosodic levels (L. Huang, personal communication, 2010; H. Lin, personal communication, 2010). Incidentally, in terms of phonemes, Inozuka et al. (2009) mentioned that Japanese is said to have around 20 phonemes, and so do European languages, including English. Thus, the number of Japanese phonemes may not necessarily be much fewer than that of English or other

languages. Here is another example of how Japanese speakers perceive the Japanese language.

d. If native Korean speakers and native Japanese speakers start to learn English at the same time, Korean learners are usually said to have better pronunciation than Japanese learners (Kindaichi, 1989).

This statement is from a textbook of the Japanese language. In order to confirm whether it is true, I asked 13 experienced ESL teachers the following question:

> Between Korean accented English and Japanese accented English, which do you think is more intelligible in general? (One book said that Korean learners of English usually had better pronunciation because Korean has more vowels than Japanese, so I would like to confirm whether it is true.)

Ten out of 13 teachers answered that Japanese ESL learners' English is generally more intelligible, and three of them answered that there is no significant difference. None of them answered that Korean ESL learners had better pronunciation. However, six of the teachers had lived in Japan for the duration of three months to 10 years, so they must be more used to Japanese accented English. Excluding them, still five of the teachers answered that Japanese accented English is more intelligible and two of them answered there is no significant difference. In more details, among the seven teachers who had never lived in Japan, one of them usually taught the beginner level, three taught the intermediate level, and three taught the advanced level of English. Two of the teachers of the intermediate level and all of the teachers of the advanced level answered that Japanese accented English is more intelligible. Therefore, the statement at issue is not supported by my informal survey. Needless to say, that Koreans' English is less intelligible does not

mean that the Korean sound system is simpler than that of Japanese. Rather, as Avery and Ehrlich (2003) mentioned that Korean ESL learners' pronunciation problems can be quite severe because of cross-linguistic sound system differences between English and Korean, the Korean sound system may happen to be more different from that of English than that of Japanese is. Although the sample size is too small to generalize, it is worthy of note that Kindaichi's (1989) perception contradicts my survey. Table 2.4 is the summary of the answers divided by the level the teachers usually taught and whether they had lived in Japan. The numbers indicate the number of the answers.

**Table 2.4**
***13 ESL teachers' answers for the question, "Between Korean accented English and Japanese accented English which do you think is more intelligible in general?"***

| Teachers | | Answers | | |
|---|---|---|---|---|
| The level they usually teach | Whether they had been in Japan | Korean accent is more intelligible. | Not different. | Japanese accent is more intelligible. |
| Beginner | Non-residents | | 1 | |
| | Residents | | | 2 |
| Intermediate | Non-residents | | 1 | 2 |
| | Residents | | | 2 |
| Advanced | Non-residents | | | 3 |
| | Residents | | 1 | 1 |

*Note*. "Beginner", "Intermediate" and "Advanced" indicate the levels the teachers usually teach. "Non-residents" indicates the teachers who had never lived in Japan. "Residents" indicates the teachers who had lived in Japan.

The following is a plan which I think would improve the textbooks. The goal is to get rid of as many unnecessarily articulatory explanations from the textbooks as possible. First, it would be better not to say anything negative, such as Japanese pronunciation is simple, or Japanese speakers are disadvantaged in acquiring pronunciation of other languages.

Simply show the English phoneme inventory and the symbol-sound correspondence rules encouraging learners not to use other sounds than those in the English phoneme inventory. Then, explain articulation of sounds that may be problematic for Japanese speakers with easy expressions. The most important point here may be to avoid claiming that sounds that are actually not problematic for Japanese speakers are problematic. Researchers and teachers would need to examine what sounds are problematic and what are not. For example, although the /s/ and /ʃ/ distinction has been widely believed to be problematic for Japanese speakers, this present study examines whether it is really problematic or not. Phonologists would need to help Japanese teachers of English and authors of pronunciation textbooks understand the sound systems of both Japanese and English more accurately by referring to results of many experiments.

To conclude, many English pronunciation textbooks in Japan tend to greatly look down upon the Japanese sound system. The attitude probably comes from the inferiority complex of the Japanese language.

## 2.11 Summary

In Japanese, [si] and [ʃi] seem to be recognized as clearly different sets of sounds by Japanese speakers. In peripheral classes of lexicon (e.g., foreign names, jargons, nicknames, names of characters in comic books and computer games), I have argued that [si] and [ʃi] are distinctive and some minimal pairs exist. With these facts, it is questionable to claim that Japanese does not have the distinction between [si] and [ʃi]. At the same time, in relatively central lexical items, I can also argue that [si] does not have a distinctive function from [ʃi], and in fact, [si] is usually replaceable with [ʃi]. In addition, there are more linguists who do not consider [si] to be distinctive from [ʃi] than linguists who do.

Therefore, that [si] is still considered to be a foreignism, or the replacement of [si] for [ʃi], implies certain connotations, such as more professional, more expressive, slangier, and so on, but it does not fully change the lexical meanings. Importantly, [si] phonetically still exists in core lexical items. In any case, the distinctive function between [si] and [ʃi] is weaker in Japanese than in English. In terms of second language acquisition, Japanese ESL/EFL learners' confusion between /s/ and /ʃ/ before high front vowels has been reported. Note that pronunciation errors do not necessarily come from L1 phonological constraints, but from other factors, such as a lack of awareness that English loanwords in Japanese and the counterparts in English are completely different. Searching for other factors than an inability to articulate is important. However, English pronunciation textbooks in Japan often attribute Japanese learners' /s/ and /ʃ/ confusion to L1 constraints or an inability to articulate. Moreover, there seems to be a tendency that authors of English pronunciation textbooks insist that the Japanese sound system is exceptionally simple among the world languages and Japanese speakers are hugely disadvantaged in acquiring foreign sounds. Because of this tendency, there is a possibility that some sounds that are actually not problematic for Japanese ESL learners are widely believed to be problematic. The /s/ and /ʃ/ distinction may be one of these. Aside from the lexically distinctive function between [si] and [ʃi], to my knowledge, there is a lack of studies that examine whether monolingual Japanese speakers, not necessarily Japanese learners of English, can distinguish between [si] and [ʃi] in both production and perception in the Japanese contexts, not in the English contexts. Therefore, this present study is designed to fill this gap. If Japanese speakers can distinguish between [si] and [ʃi], pronunciation teaching materials would have to consider a different approach regarding Japanese ESL learners' [s] and [ʃ] confusion from

the current approach where Japanese speakers are considered not to have this distinction in their L1. Thus, this present study is important in the area of second language acquisition.

# Chapter Three

# EXPERIMENT A: [si] AND [ʃi] IN JAPANESE CONTEXT

## 3.1 Research Design

Can monolingual Japanese speakers distinguish between [si] and [ʃi] in Japanese in both production and perception? Previous studies are mostly about how Japanese learners of English perceive [si] and [ʃi] in the *English* contexts. In the course of this research, I examined how Japanese speakers, not necessarily learners of English, perceive and produce [si] and [ʃi] in the *Japanese* contexts. In other words, it is about to what extent the syllable [si] is recognized in the Japanese language. It is often said that seniors cannot pronounce some syllables used only in foreign originated words, including [si], but as far as I looked for, I could not find any research regarding how older generations pronounce [si]. I designed three experiments, two for production and one perception.

- Experiment 1: Participants read aloud the written stimulus without hearing sound cues.
- Experiment 2: Participants repeated the sound stimulus without seeing written cues.
- Experiment 3: Participants listened to the sound stimuli recorded by a phonetically trained native speaker of Canadian English and wrote down what they heard.

The point of these experiments is to not let participants associate the stimuli with English in order to reduce psychological factors that could cause participants' errors, such as over-generalization.

## 3.2 Experiment 1: Reading Aloud

### *3.2.1 Objective*

This experiment examines whether Japanese speakers can distinguish between [si] and [ʃi] in production when they see a nonsense sentence written in the Japanese

orthography. In this experiment, they did not hear a sound stimulus in order to draw out their production without a bias toward a sound stimulus.

### *3.2.2 Participants*

I organized two groups, a group of monolingual Japanese speakers (JG1) and a group of Japanese ESL learners (JG2). JG1 consisted of 93 monolingual native speakers of Japanese who lived in Tôkyô or the adjacent prefectures at various ages in the range of 17 to 89. Most of them had never lived in an English-speaking country. One participant who had been most exposed in a foreign language was a female in her 20's who lived in France for about half a year but not in an English-speaking country. Another two females in their 20's had been in British Columbia for about two months but were not exposed in an English environment very much and could not pronounce most of the English phonemes, such as /l/, /ɹ/, /θ/, /ð/, and /v/, that did not exist in Japanese. JG2 consists of 12 Japanese speakers who lived in Japan but had some exposure to English, with the exception of four participants, one male in his 30's, one female in her 30's, one female in her 20's and one female in her teens, who lived in British Columbia. In other words, they were ESL learners. Their ages were in the range between 18 and 78. The participant who was exposed the least to English was a female in her 30's who had been in English-speaking countries for a little more than two months in total. The most exposed participant was a female in her 30's who had been in British Columbia for about five years but lived in Tôkyô when the experiment was performed. Table 3.1 is the number of participants in each age group and gender in JG1 and JG2.

**Table 3.1**
***The number of participants in each group of age and gender in JG1 and JG2***

| | JG1 | | | JG2 | | |
|---|---|---|---|---|---|---|
| Age | Male | Female | Total | Male | Female | Total |
| 10-19 | | 1 | 1 | | 1 | 1 |
| 20-29 | 6 | 18 | 24 | 1 | 2 | 3 |
| 30-39 | 4 | 6 | 10 | 1 | 5 | 6 |
| 40-49 | 6 | 8 | 14 | | 1 | 1 |
| 50-59 | 4 | 10 | 14 | | | |
| 60-69 | 15 | 6 | 21 | | | |
| 70-79 | 5 | 1 | 6 | 1 | | 1 |
| 80-89 | 2 | 1 | 3 | | | |
| Total | 42 | 51 | 93 | 3 | 9 | 12 |

*Note.* JG1 = monolingual native speakers of Japanese; JG2 = Japanese ESL learners

### *3.2.3 Written Stimulus*

みずはしバルスィ　シーセルズ　スィーシェルズ　バイザ　スィーショール

/midzɯhaʃibaɾɯsi(:)　ʃi:seɾɯdzɯ　si:ʃeɾɯdzɯ　baidza　si:ʃo:ɾɯ/[24]

The first part derives from the name of a character in a computer game, /midzɯhaʃi paɾɯsi:/, in which both [ʃi] and [si] occur, as mentioned in 2.4. Since this word is used only by a certain young clientele of the computer game, most of the participants did not seem to know it. In the stimulus, I changed /p/ in the original word into /b/ in order to avoid vowel devoicing; /i/ before /p/ is likely to be devoiced due to the allophonic alternation that high vowels between voiceless obstruents are devoiced in standard Japanese. The rest of it is from the English tongue twister "She sells seashells by the seashore." I changed it into my

[24] Again, in this paper, I sometimes use the phonemic bracket / / for /si/ and /ʃi/ in Japanese in order to clarify the difference between underlying representations and surface representations because in surface representations, some phonetic variations appeared of both /si/ and /ʃi/.

version of *Katakana*-English.[25] Since *Katakana* (the Japanese syllabary) has no way to distinguish between /ð/ and /z/, and /ɹ/ and /l/, those distinctions disappeared, but the most important distinction at issue, /s/ versus /ʃ/, remained. The reason I chose it was that as mentioned in 1.1, Fukushima, O'Connor, and Matsuzawa (2008) found that about 2,500 out of 3,000 Japanese learners of English could not correctly distinguish between /s/ and /ʃ/ in this English tongue twister, and I wondered if Japanese speakers can distinguish those consonants when this tongue twister is written with the Japanese orthography. Therefore, I had the participants believe that this stimulus was completely Japanese. Fortunately, most of the participants in JGs did not know this tongue twister. Consequently, the whole part of this stimulus looked just like a set of nonsense words. In cases when a few of the participants recognized that it was from the English tongue twister, I urged them not to associate the stimulus with it.

### *3.2.4 Procedure*

The participants read the sentence out loud twice and their reading was recorded with an electrets condenser microphone, SONY ECM-MS908C using Praat set at 44100Hz, and each recording was saved as a WAV file. Recording was done in a quiet room. I let the participants practice a few times before recording if they wanted to. The judges picked the better one of the two trials if one of the trials was incorrect or less clear.

---

[25] The term "*Katakana*-English" is used in various aspects, such as phonology, morphology, and semantics. In this paper, I use this term in the sense of pronunciation, and I define "*Katakana*-English" as English commonly pronounced by Japanese EFL learners in the completely Japanese phonological way. In my definition, *Katakana*-English words are not necessarily a part of the Japanese vocabulary. Since *Katakana*-English pronunciation often have gone through complicated loanword adaptation processes, its pronunciation often may not be predictable from the original English pronunciation. *Katakana*-English pronunciation also may vary from individual to individual.

### *3.2.5 Rating*

There were three Japanese and one American judges. The Japanese judges had different linguistic backgrounds: a male monolingual Japanese speaker in his 70's, a female Japanese speaker in her 30's who had stayed in English-speaking countries for a few months, and me. The American judge was a phonetically trained female in her 20's. The three Japanese judges listened to the recordings and discussed them to make the final decision. Three of us perceptually judged if the participants properly distinguished between [si] and [ʃi], as well as if their [si] had a detectible on-glide, like [swi] or [sɯi].[26] When a sound was perceptually difficult to judge, I visually checked the spectrogram on Praat. After that, the American judge and I listened to the recordings and discussed them to make the revised final decision. The American judge perceptually judged if the participants' [s] and [ʃ] were categorized as the English /s/ and /ʃ/ (Eng/s/ and Eng/ʃ/) by English ears. We also judged if some of their [ʃ] "could" be perceived as Eng/s/ and some of their [s] "could" be Eng/sw/ or Eng/θ/ by English ears. When facing difficult sounds to transcribe, I asked another English-speaking phonetician. We stayed with the "could"-assumption because we cannot draw a conclusion unless performing a perception test with a number of native English speakers, and because there are phonetically no clear-cut segment boundaries especially around the alveolar region.

### *3.2.6 Results and Discussion*

All participants in both JG1 and JG2 clearly distinguished between /ʃi/ and /si/ regardless of the participant's age. Recall that /ʃi/ occurred twice and /si/ occurred three times in the stimulus. Only one male in his 40's in JG1 and one female in her 30's in JG2

---

[26] The symbol [ɯ] is the velar approximant, and [w] is the labial-velar approximant.

pronounced one occasion of /si/ wrongly, but since all the other occasions were correctly pronounced, the judges concluded that those participants simply made an accidental mistake. This result contradicts the general belief that Japanese speakers cannot distinguish between /ʃi/ and /si/ or that older speakers cannot distinguish between them. Since the stimulus was a nonsense sentence and originated from a tongue twister, some participants found it somewhat difficult, but still many of them pronounced it easily. This indicates that Japanese speakers consider /ʃi/ and /si/ as two different entities. What is interesting is that in the observation by Fukushima et al. (2008), over 80% of their participants mixed up /ʃi/ and /si/ in the English tongue twister, while in contrast, in my experiment, 100% of my participants correctly distinguished between /ʃi/ and /si/ in the same tongue twister. In other words, Japanese speakers can distinguish between those syllables in Japanese, but when it comes to the English contexts, they somehow get confused. An experiment has been designed to examine this discrepancy, which is introduced in Chapter Four.

Meanwhile, the quality of /si/ is very intriguing. In Experiment 1, 39 of the participants in JG1 and five in JG2 tended to pronounce /si/ as [sɯi] or [swi] with on-glide. For example, Figure 3.1 is the spectrogram of /si/ and /ʃi/ up to 5000 Hz produced by a male in his 40's who always pronounced on-glide. In the nucleus part of the /si/ syllable, F2 clearly goes up, which means the tongue moved from back to front. In contrast, the /ʃi/ syllable shows no glide. Figure 3.2 is the spectrogram of /si/ and /ʃi/ produced by a male in his 20's who always pronounced pure [si] without glide. In the nucleus of his /si/ syllable, F2 is in the steady state. Incidentally, [ʃ] has high amplitude in the lower frequency region than [s] does. This indicates that these speakers distinguished [s] and [ʃ] in quality.

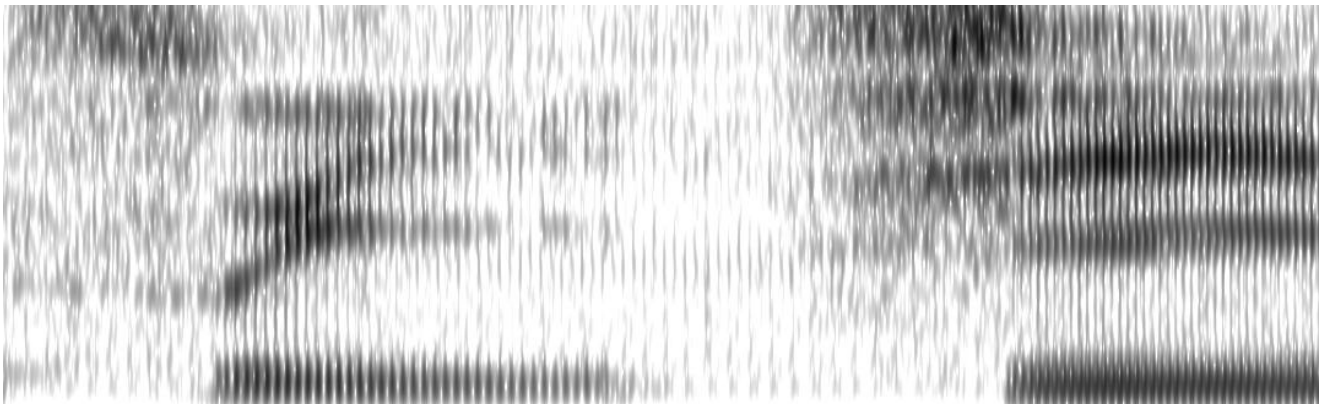


***Figure 3.1.*** **Clear on-glide after [s] produced by a male in his 40's**

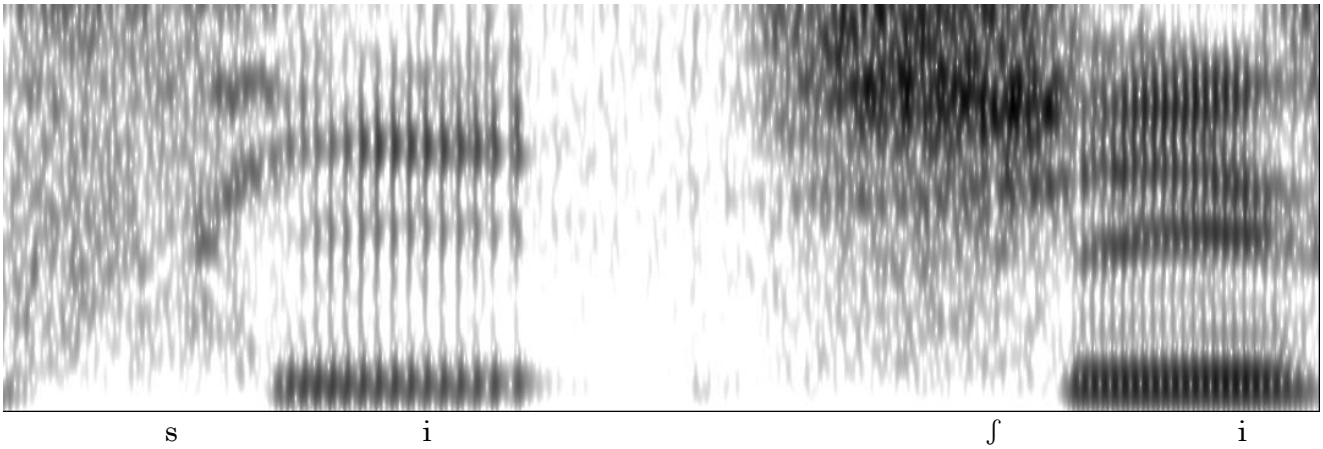


***Figure 3.2:*** **[s] not followed by on-glide produced by a male in his 20's**

## 3.3 Experiment 2: Mimicking

### *3.3.1 Objective*

In Experiment 1, participants might have been influenced by the orthography. Recall that there are some variations in the ways of writing foreign originated sounds with the Japanese orthography as mentioned in 2.3. It is possible that the orthography "スィ" in the stimulus in Experiment 1 was associated with the sound [si] for some people but was associated with a different sound for others. To eliminate the orthographical influence, Experiment 2 was designed with exclusive auditory stimulus.

### *3.3.2 Participants*

The participants were the same as those who participated in Experiment 1.

### *3.3.3 Sound Stimulus*

/ʃiNgaː no joʃʃiː/ and /siNgaː no jossiː/

The first one means "A singer called 'Yosshî.'" The second one means "A singer called 'Yossî.'" Incidentally, both /ʃiNgaː/ and /siNgaː/ mean "singer," except [si] sounds more fancy. The stimulus was pronounced by a monolingual Japanese female in her late 20's who had received training on her Japanese pronunciation in announcer training courses.[27] She was recorded reading the stimulus and instruction with an electrets condenser microphone, SONY ECM-MS908C, with Praat set at 44100Hz and saved as a WAV file.

### *3.3.4 Procedure*

The participants listened to the stimulus and the instruction with headphones, Pioneer SE-M390 (Frequency response: 5~29000Hz). After that, I orally demonstrated the stimulus in order to make sure the participants understood it. Then, I let the participants practice before recording if they wanted to. The participants were recorded pronouncing it twice just as they did in Experiment 1. The recordings were judged in the same way as Experiment 1. When the data were analyzed, correlation test was used to examine correlations among variables.

### *3.3.5 Rating*

The rating was done in the same way by the same judges as in Experiment 1.

---

[27] She also explained, "The first one has a purely Japanese-like syllable [ʃi], while the second one has a little fancy syllable [si]."

### *3.3.6 Results and Discussion*

The results show once again that all of the participants distinguished [si] and [ʃi].[28] The two participants who made an accidental mistake in Experiment 1 performed correctly in Experiment 2. This time, many of them repeated the stimulus correctly with no difficulty, while a few of them, especially participants who were 70 years old and older, had trouble in memorizing it or were somewhat confused by this tongue twister-like stimulus. Some of them even commented that they were visual learners and were not good at listening. Three participants, two males in their 60's and one female in her 80's, in JG1 pronounced [siNga: no joʃʃi:] which was supposed to be [siNga: no jossi:]. However, they still produced both [si] and [ʃi]. This implies some possibilities. Firstly, reading materials can be easier than listening materials since in Experiment 2, the participants were required to not only perceive and produce, but also memorize. Secondly, according to Masuda's (2003) study on perception of English consonants by Japanese EFL learners mentioned in 2.7, when the target sound appeared word-initially, listeners had more chance to get correct answers.

In Experiment 2, 31 of the participants in JG1 and three in JG2 tended to pronounce /si/ as [sɯi] or [swi] with on-glide. More interestingly, some individuals seemed to purposely distinguish between pure [si] and the one with on-glide, which is discussed next.

### *3.3.7 Discussion of Experiments 1 and 2: /si/ or /swi/*

We judged the quality of [si] based on how native English speakers would categorize it. As mentioned above, the judges just assumed whether native English speakers "could" categorize it as the Eng/si/ or Eng/swi/. It is because the judges cannot draw a conclusion

---

[28] In JG1, two male participants in their 20's had Down's syndrome, but they still distinguished between [si] and [ʃi]. This can emphasize that [si] and [ʃi] are recognized differently by Japanese speakers.

regarding the sound qualities unless perception tests with native English speakers are conducted.

As for [si], 47 in JG1 and six in JG2 did not add detectable on-glide in both Experiment 1 and 2; in other words, their [si] was categorized as Eng/si/. The other 46 in JG1 and six in JG2 tended to pronounce [sɯi] or [swi] with on-glide, and occasionally even almost two-mora-like [sɯi], which can be perceived as /swi/ or even /sui/ by native speakers of English. The production of half of the participants had this characteristic. Recall that TuZino (2008) stated that in the modern writing system, "スィ" can be pronounced as either [si] or [sʷi], although [si] or [sʷi] are not phonologically contrastive, as mentioned in 2.3. Moreover, people who added on-glide can be divided into four types:

1. In the first type in which 19 of them were grouped, they quite randomly pronounced either [si] or [swi]. Interestingly, when the target was a one-mora syllable /si/, they tended not to add on-glide while when the target was two-mora syllable /si:/, they tended to add on-glide regardless of whether it was in Experiment 1 or 2.
2. In the second type in which 15 were grouped, they tended to pronounce [swi] in Experiment 1 (written stimulus) while they produced pure [si] in Experiment 2 (sound stimulus). This asymmetry may be due to the orthographical influence. As explained in 2.2 and shown in Table 2.2, in the Japanese orthographical rules, a small *Hiragana* or *Katakana* symbol (Japanese syllabary) is not pronounced or counted as one mora. Therefore, "ィ" in "スィ" is not supposed to be counted as a mora; "スィ" should be pronounced as [si] within one mora. However, as Matsuzaki (1992) mentioned, in some particular foreign originated words in Japanese, people sometimes over-generalize the rule that one symbol corresponded to one mora, and

consequently, they mix up a non-moraic small symbol with a one-mora large symbol. In this case, those participants must have almost mixed up the small "ィ" with the one-mora large "イ" with a dilemma between the small one and the large one, and ended up somewhere between no mora and one mora. In other words, they mixed up "スィ" [si] with "スイ" [sɯi] and ended up somewhere in between. In contrast, when they were asked to simply imitate what they heard, they pronounced [si] without being distracted by the orthography. That indicates that they actually own the [si] syllable in their inventory but get easily distracted.

3. The third type in which seven participants were grouped shows the opposite pattern to the second group. They did not pronounce on-glide in Experiment 1 (written stimulus) but did so in Experiment 2 (sound stimulus). Since in the Experiment 2 they were urged to make a clear difference between [ʃi] and [si], they seemed to overdo it. Or, because the narrator mentioned [ʃi] as a pure-Japanese-like way and [si] as a "fancy" way, the participants tried to make a clear contrast by making a Japanese-like sound and a fancy sound, instead of repeating what they actually heard. In any case, they own the [si] sequence in their inventory but easily lean to [sɯi] or [swi] when distracted by other factors, such as personal preferences.
4. In the fourth type in which five were grouped, they always quite clearly added on-glide and never pronounced pure [si]. A question here is whether they substituted [si] with the two-mora [sɯi]. Figure 3.3 shows the /si:ʃ/ part from Experiment 1 produced by a male in his 60's. In this phonetic context, the boundaries between the consonants and vowels are clear. Since in Japanese two-mora vowels are usually almost twice longer than one-mora vowels (Hirata, 2004), there should be [ɯ] that is

half the length of [i:] if he pronounced /si:ʃ/ as [sɯi:ʃ]. However, Figure 3.3 shows the categorical divisions based on the perceptual impression, in which the velarized vocalic part is too short to be considered moraic. Based on the visual cues, the steady state of F2 is more than three times longer than the curvy part. It is considered as the glide [ɯ̯], rather than the moraic [ɯ]. Some of the participants occasionally pronounced two-mora-like [sɯi], but in most of the cases they pronounced the glide [ɯ̯]. Therefore, I conclude that they consider /si/ as a different entity from /sɯi/.

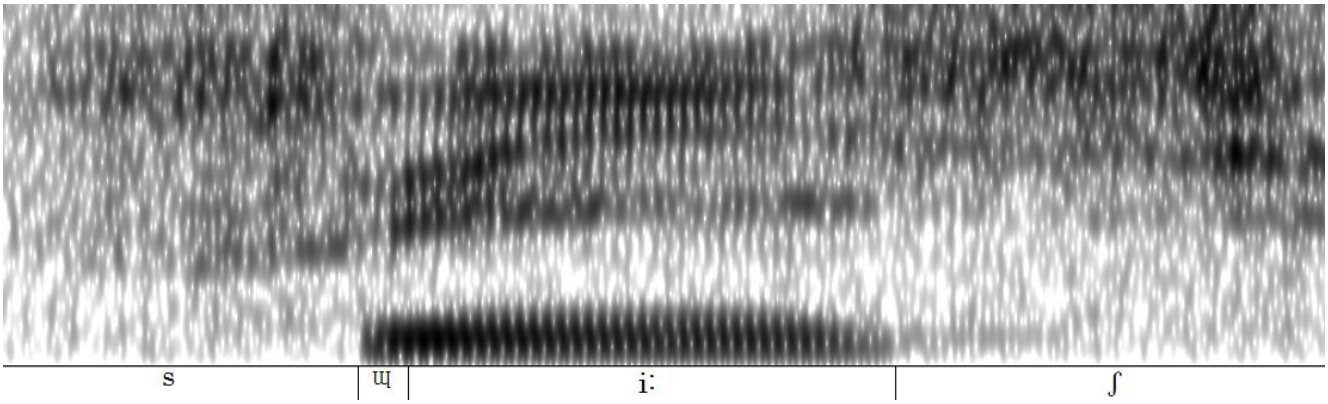


***Figure 3.3.*** **/si:ʃ/ produced by a male in his 60's**

The next question is where this on-glide comes from. For example, according to Matsuzaki's (1993) paper, the syllables /ɸa/, /ɸi/, /ɸe/, and /ɸo/ were already considered to be phonotactically allowed in Japanese by most linguists. As well, in the writing system, these syllables were included in the set of official syllables by the Ministry as mentioned with Figure 2.1. It is not unusual that a two-mora sound eventually becomes a one-mora sound when a foreign sound is mixed in Japanese. For instance, /**ɸi**ɾɯmɯ/ (film) and /**ɸe**ɾɯto/ (felt) were pronounced as /**ɸɯi**ɾɯmɯ/ and /**ɸɯe**ɾɯto/ respectively by older generations (Koizumi, 2003). These are the examples of two-mora (/ɸɯi/ and /ɸɯe/) becoming one-mora (/ɸi/ and /ɸe/). Matsuzaki (1992) showed some words that have a sound varying from one mora to

two moras. One of the examples is “whi” in “whisky” which was imported to Japanese and pronounced as either [wi (ɯi)] (one mora) or [ɯi] (two moras). Another interesting example from Matsuzaki’s (1992) paper is “Sweden” where “Swe” can be pronounced as either [sɯwe: (sɯɯe:)] or [sɯe:] in Japanese and there are infinite numbers of steps between [ɯwe: (ɯɯe:)] and [ɯe:] depending on speakers and occasions. Note that [wi] and [we] have not yet been unanimously admitted to be in the Japanese phonological system (Matsuzaki, 1993).

Considering this tendency of the Japanese language, it can be assumed that the status of /si/ in Japanese is on the way from [sɯi] (two mora) to [si] (one mora), as well as the fact that there are infinite steps between [si] and [sɯi] depending on speakers and situations. The result that half of the participants pronounced pure [si] implies that about 50% of them have completed the 1-mora-to-2-mora process while the other 50% are still in the process.

### *3.3.8 Discussion of Experiments 1 and 2: /s/ or /θ/, /ʃ/, or /s/*

We judged whether native English speakers “could” categorize their [s] and [ʃ] as Eng/s/, Eng/θ/, or Eng/ʃ/.

The sound [s] produced by seven of the participants sometimes could be categorized as the English /θ/. However, it was not completely Eng/θ/, but difficult to be decided as /s/ or /θ/. As well, [θ]-like [s] only sometimes appears and the frequency of occurrence varies from speaker to speaker. What is interesting is that despite the general belief that young people tend to pronounce [θ]-like [s], in my experiment, participants in their 60’s did it as well. As the assumption of Inozuka, et al. (2009) mentioned in 2.6, this pronunciation is probably a type of weakened version of [s]. In this paper, the judgement is based only on whether native English speakers could categorize it as Eng/θ/; therefore, actually more

participants' [s] might be different from the prototypical Japanese [s], but still categorized as the Eng/s/ by native English speakers. Probably, [s] produced by the Japanese is not likely to be completely categorized as Eng/θ/ by English ears, but further perception tests are required to verify my findings.

As for [ʃ], in Experiment 1 (written stimulus), it can be categorized as Eng/ʃ/ in most of the cases. However, the word-initial [ʃ] in Experiment 2 produced by 24 participants in JG1 and three in JG2 is difficult to be judged as either Eng/ʃ/ or Eng/s/ unless they are compared with the [s]s produced by the same speaker. Figure 3.4 is the spectrogram up to 8000Hz from one of the 24 participants in JG1 who pronounced [s]-like [ʃ] in word-initial but the clear [ʃ] in the other environments in Experiment 2.

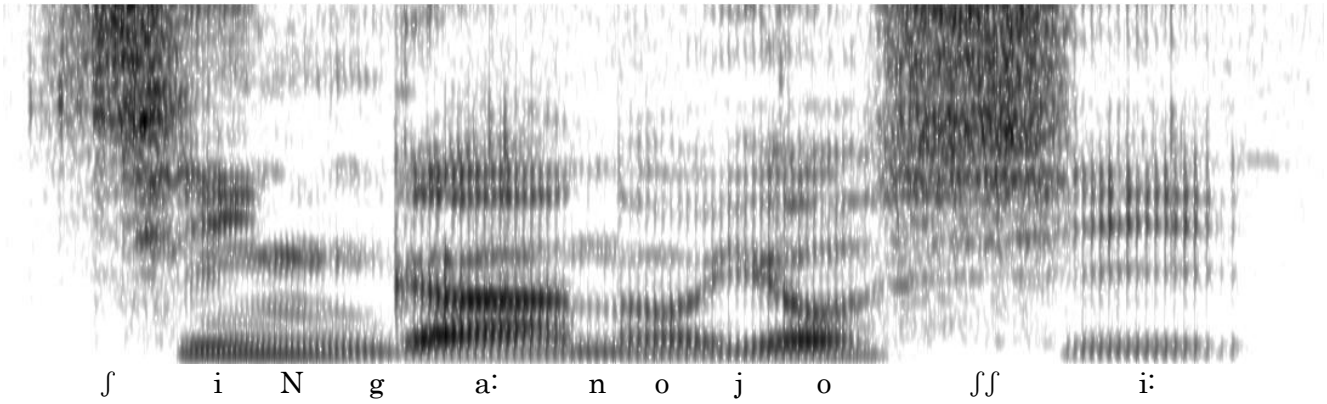


***Figure 3.4.*** **Different qualities of [ʃ] at word-initial and word-final produced by a male in his 40's**

In the first /ʃ/, the energy does not quite spread to below 4000Hz. As mentioned in 2.6, the energy of the Japanese /s/ (Jap/s/) concentrates around 4000Hz while that of Jap/ʃ/ spreads to the lower region; therefore, the beginning part sounds like [s] and it is gradually getting closer to [ʃ], which can be described as [sʃ]. In contrast, the word-medial /ʃʃ/[29] part shows steady frication. Actually, many of the participants, including those whose [ʃ] can still be

[29] In Japanese, a single consonant and a geminate consonant are phonemically distinct.

categorized as Eng/ʃ/ but not Eng/s/, showed a similar pattern. In detail, 19 of the 24 participants in JG1 and all of the three participants in JG2, who pronounced [s]-like [ʃ] at word-initial in Experiment 2, actually pronounced [sʃ] that could be categorized as Eng/s/ by English ears. Interestingly, 13 of those 19 participants in JG1 and all of those three in JG2 did not clearly show this characteristic in Experiment 1. One of the possible reasons is that in the stimulus in Experiment 1, [ʃ] does not appear sentence-initially and it was pronounced inter-vocalically unless the participants read the stimulus in a choppy way. Another possible reason is that written cues may have encouraged the participants to pay more attention to each sound. The rest of the six participants out of the 19 also pronounced [sʃ] in the second phrase in Experiment 1 although it is still closer to Eng/ʃ/. In the stimulus "みずはしバルスィ　シーセルズ　スィーシェルズ　バイザ　スィーショール," phonologically transcribed as /midzɯhaʃibaɾɯsi(:) ʃi:seɾɯdzɯ si:ʃeɾɯdzɯ baidza si:ʃo:ɾɯ/, in which the second phrase "シーセルズ" /ʃi:seɾɯdzɯ/ starts with /ʃ/, many of the participants made a pause before it, and this /ʃ/ can be in the same environment as that of the word-initial one in Experiment 2. Therefore, there is a tendency that those participants pronounce post-pausal [ʃ] as [sʃ]. The other five participants of the 24 in JG1 whose [ʃ] can be heard as Eng/s/ did not clearly show [sʃ] in the spectrogram, but the energy of their word-initial [ʃ] spreads relatively high in frequency compared to the other occasions of [ʃ] within the same speaker. The question is why they pronounced [sʃ] or [s]-like [ʃ] in pre-pausal position. It can be because their tongue moves from the neutral position to the front part of the hard palate. Those who pronounced [sʃ] started to make frication before their tongues got close enough to the hard palate. Although the word-initial [ʃ] in the second phrase of the stimulus in Experiment 1 is after a pause (midzɯhaʃibaɾɯsi(:) ʃi:seɾɯdzɯ), the

first phrase finishes in [i(ː)] and their tongues were still close enough to the palate, which made them less likely to pronounce [sʃ].

Another intriguing finding is that the word-initial [ʃ] in Experiment 2 (sound stimuli) produced by the last three participants, two in their 20's and one in her teens in JG1, are considered highly likely to be categorized as Eng/s/ by English ears, although their /ʃ/ was phonetically different from the same person's /s/. For example, Figure 3.5 is the comparison between post-pausal [ʃi] and post-pausal [si] produced by a female in her late 20's. The spectrogram is shown up to 6000Hz and there are spectral slices of 50 milliseconds (msec)[30] in the steady state portion of both [ʃ] and [s]. The spectral slices were smoothed by LPC smoothing set at five peaks.

[30] According to Pan et al. (2004), there is a way to pick up 100msec of the fricative part, as Ladefoged did (2003), but Pan et al. took only 50msec in case spectral patterns change within the fricative. I followed Pan at al.'s way because the spectral patterns in the fricative do not stay steady for 100msec in this case.

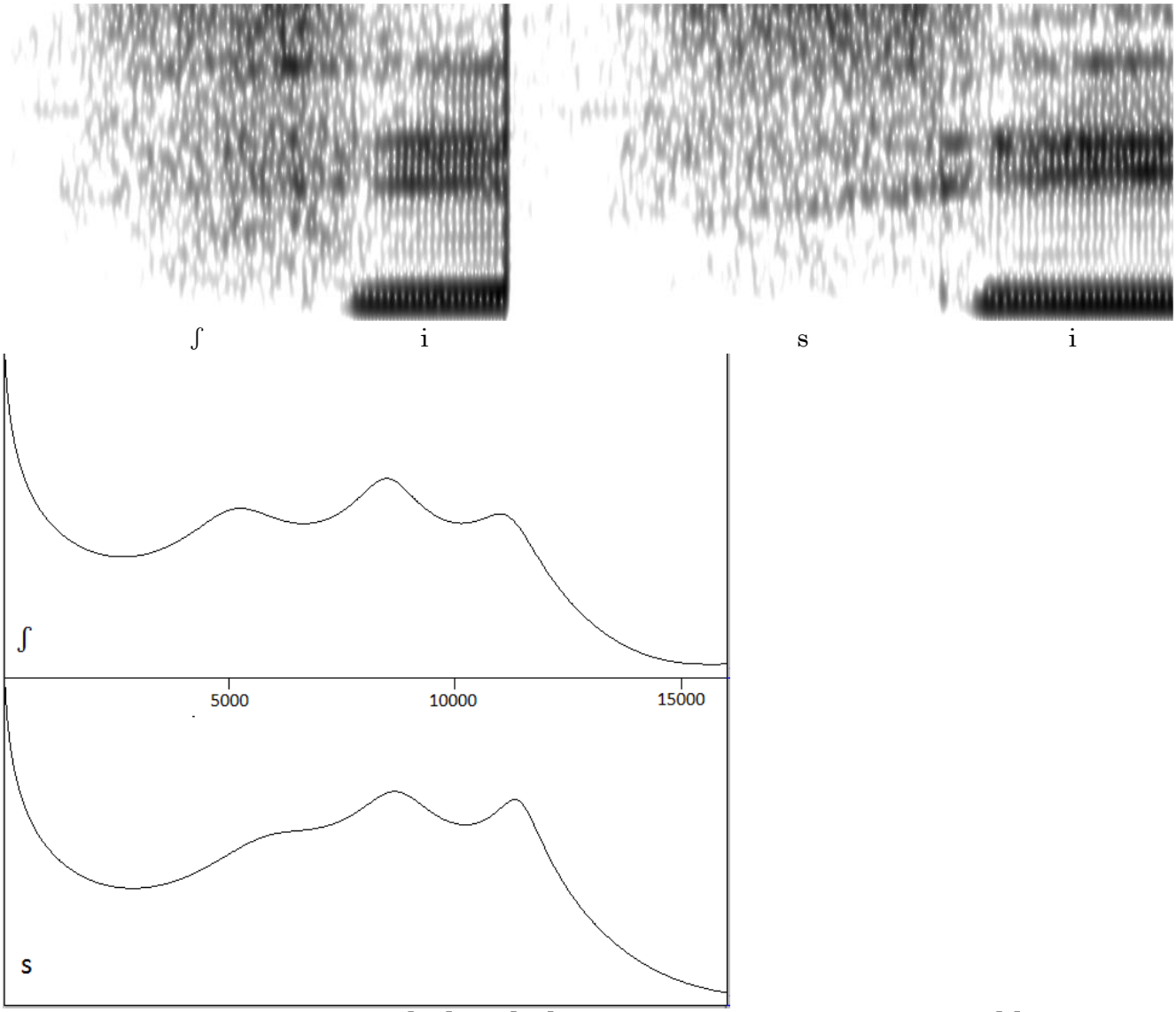


***Figure 3.5.*** **The comparison between [#ʃi] and [#si] produced by a female in her 20's whose [ʃ] can be categorized as Eng/s/**

Beckman et al. (2009) mentioned that Jap/ʃ/ and Jap/s/ can be separated in the centroid (centre of mass of the distribution) and the onset of F2 of the following vowel. The centroid of this participant's /ʃ/ is 7349Hz while that of her /s/ is 8514Hz calculated by Praat, which may not be clearly different enough. However, Figure 3.5 shows a clearer difference; the first peak in her /ʃ/ appears around 5200Hz while the first peak in her /s/ appears around 8600Hz. As well, F2 onsets of the vowels following her /ʃ/ and /s/ were measured at the

closest 0 crossing from the first pulse by Praat. The onset of the vowel after /ʃ/ shows 2791Hz and the middle of the vowel shows 2633Hz, while the onset of the vowel after /s/ shows 2550 and the middle of the vowel is 2681Hz. /ʃ/ shows higher F2 onset as Beckman et al. mentioned. Since both centroid and F2 onset show different patterns, the speaker made a contrast even if both fricatives can be categorised as Eng/s/ by English ears. It could be "covert contrast[31]" for English ears. For Japanese ears, her [ʃ] and [s] could still sound phonologically distinctive based on my perceptual impression as a native Japanese speaker, although I cannot draw a conclusion unless I perform perception tests. Other occasions of [ʃ] pronounced by these three participants, whose /ʃ/ is close to Eng/s/, are also sometimes more like Eng/s/ than Eng/ʃ/, especially word-initially in Experiment 2. As it is generally said that young Japanese speakers' pronunciation of Jap/ʃ/ is not clearly different from Jap/s/, these three are good examples. Just like Inozuka et al.'s (2009) assumption that [θ]-like [s] is probably a type of weakened version of [s], [s]-like [ʃ] is perhaps a weakened version of [ʃ] as well. At this moment, it is not clear whether the quality of their /ʃ/ is an ESL/EFL issue or an issue within the Japanese language. Table 3.2 and 3.3 is the summary of the quality of their [s] and [ʃ]. The "target" column is the target sound in the stimuli in both Experiment 1 and Experiment 2. "Production" is what the participants produced being judged based on what native English speakers could perceive. In Table 3.2, all the participants were put into either the [si] group or the [swi] group. The [swi] group is divided into four groups: 1) those who randomly pronounced [w], 2) those who pronounced [w] only in Experiment 1 (written stimulus), 3) those who pronounced [w] only in

---

[31] Covert contrast is an acoustically significant difference, which is not large enough to be perceived reliably (Beckman, Edwards, & Yoneyama, 2003).

Experiment 2 (sound stimulus), and 4) those who always pronounced [w].

**Table 3.2**
***The summary of the quality of glide insertion***

| | JG1 Target /si/ | | | | | JG2 Target /si/ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Production | | [si] | [swi] | | | | [si] | [swi] | | | |
| Age (Gender) | | | *Ran-dom* | *Only Ex1* | *Only Ex2* | *Al-ways* | | | *Ran-dom* | *Only Ex1* | *Only Ex2* | *Al-ways* |
| 10-19 (M) | | | | | | | | | | | | |
| (F) | N=1 | | 1 | | | | N=1 | 1 | | | | |
| 20-29 (M) | N=6 | 3 | 1 | 1 | 1 | | N=1 | | | | | 1 |
| (F) | N=18 | 14 | 2 | 1 | 1 | | N=2 | 2 | | | | |
| 30-39 (M) | N=4 | 2 | | 1 | 1 | | N=1 | | | 1 | | |
| (F) | N=6 | 6 | | | | | N=5 | 2 | 1 | 1 | 1 | |
| 40-49 (M) | N=6 | 2 | 1 | | 1 | 2 | | | | | | |
| (F) | N=8 | 4 | 1 | 3 | | | N=1 | 1 | | | | |
| 50-59 (M) | N=4 | 1 | | 1 | 2 | | | | | | | |
| (F) | N=10 | 6 | 2 | 2 | | | | | | | | |
| 60-69 (M) | N=15 | 3 | 8 | 1 | | 3 | | | | | | |
| (F) | N=6 | 2 | 1 | 2 | 1 | | | | | | | |
| 70-79 (M) | N=5 | 2 | 2 | 1 | | | N=1 | | | 1 | | |
| (F) | N=1 | 1 | | | | | | | | | | |
| 80-89 (M) | N=2 | 1 | | 1 | | | | | | | | |
| (F) | N=1 | | | 1 | | | | | | | | |
| | | | 19 | 15 | 7 | 5 | | | 1 | 3 | 1 | 1 |
| Total | N=93 | 47 | 46 | | | | N=12 | 6 | 6 | | | |

*Note*. All the participants are put in one of the categories. N = the number of the participants in each age and gender group; (M) = male; (F) = female

Table 3.3 shows only the numbers of 1) those who sometimes pronounced [θ]-like [s] (indicated as "θi – si"), 2) those who pronounced [ʃ] which was highly likely to be heard as Eng/s/ at word-initial and as also a sound between [s] and [ʃ] at other phonetic environments (indicated as "almost si"), and 3) those who pronounced [sʃ] or [s]-like [ʃ] predominantly at word-initial in Experiment 2 (indicated as "si – ʃi *wrd-int in Ex2*"). Those whose /ʃ/ can be clearly categorized as Eng/ʃ/ are not counted in Table 3.3.

**Table 3.3**
***The summary of the quality of the participants' [s] and [ʃ]***

| | JG1 | | | | JG2 | | | |
|---|---|---|---|---|---|---|---|---|
| Target | /si/ | /ʃi/ | /ʃi/ | | /si/ | /ʃi/ | /ʃi/ | |
| Production | [θi] – [si] | *almost* [si] | [si] – [ʃi] *wrd-int in Ex2* | | [θi] – [si] | *almost* [si] | [si] – [ʃi] *wrd-int in Ex2* | |
| Age (Gender) | | | | | | | | |
| 10-19 (M) | | | | | | | | |
| (F) N=1 | | 1 | | | N=1 | | | |
| 20-29 (M) N=6 | 1 | | 1 | | N=1 | | | |
| (F) N=18 | | 2 | 10 | | N=2 | | | |
| 30-39 (M) N=4 | | | | | N=1 | | | |
| (F) N=6 | 1 | | 3 | | N=5 | | 3 | |
| 40-49 (M) N=6 | | | 1 | | | | | |
| (F) N=8 | | | 1 | | N=1 | | | |
| 50-59 (M) N=4 | | | | | | | | |
| (F) N=10 | | | 1 | | | | | |
| 60-69 (M) N=15 | 3 | | 4 | | | | | |
| (F) N=6 | 2 | | 1 | | | | | |
| 70-79 (M) N=5 | | | 1 | | N=1 | | | |
| (F) N=1 | | | 1 | | | | | |
| 80-89 (M) N=2 | | | | | | | | |
| (F) N=1 | | | | | | | | |
| | | 3 | 24 | | | | 0 | 3 |
| Total N=93 | 7 | 27 | | | N=12 | 0 | 3 | |

*Note*. Only the participants with these characteristics are put in this chart. N = the number of the participants in each age and gender group; (M) = male; (F) = female

If it is assumed that the /si/ syllable in the Japanese language is in the process of development from [si] to [swi], there could be a correlation between age and presence of on-glide. 47 participants in JG1 and six participants in JG2 pronounced [si], while 46 in JG1 and six in JG2 added on-glide regardless. The correlation between the two pronunciation groups, the "si" group and the "swi" group, and eight age groups in JG1 is $r = 0.29$ ($p < 0.005$). The r value was calculated in Pearson by inputting the age group (10, 20, 30...) and the pronunciation type (1 = si, 2 = swi) for each of 93 participants. If JG1 and JG2 are combined, the correlation is still $r = 0.29$ ($p < 0.005$). There is a weak correlation. Some participants in their 10's and 20's had not completed the process from [swi] to [si] yet,

whereas some in their 60's, 70's and 80's had already completed it. This sound shift seems to be going so slowly that there is no clear-cut boundary between "younger generations" and "older generations." Another interesting point is the gender difference; more females pronounced pure [si] than males. The correlation between gender and the pronunciation type is $r = 0.31$ ($p < 0.005$). The variables are gender (1 = female, 2 = male) and the pronunciation type (1 = si, 2 = swi). There is a moderate correlation. To sum up, more older people insert glide than younger people, and more males insert glide than females.

There is no significant correlation between the age groups and whether the participants pronounced [s] or [θ]-like [s] ($p > 0.1$). The variables are the age group (10, 20, 30...) and the pronunciation type (1 = [θ]-like [s], 2 = [s]). That in the combination of JG1 and JG2 is not significant, either ($p > 0.1$). Five of the seven participants who pronounced [θ]-like [s] pronounced [swi]. Therefore, perhaps those who pronounce pure [si] tend not to pronounce [θ]-like [s].

Interestingly, the correlation between the generation groups and whether they pronounced [ʃ], somewhat [s]-like [ʃ] at word-initial or quite [s]-like [ʃ] is $r = 0.34$ in JG1 ($p < 0.001$). The variables are the age group and the pronunciation type (1 = more [s]-like [ʃ], 2 = somewhat [s]-like [ʃ] at word-initial, 3 = [ʃ]). The correlation is moderate. In fact, all of the three participants who pronounced quite [s]-like [ʃ] were in their teens or 20's. As well, 11 of the 24 participants in JG1 who pronounced somewhat [s]-like [ʃ] at word-initial were in their 20's. In particular, young females seem to have this tendency. If JG1 and JG2 are combined, the correlation slightly drops to $r = 0.30$ ($p < 0.005$). They might have been influenced by the English post-alveolar [ʃ] to some extent, but at this moment, I cannot conclude if they easily change the quality of [ʃ] from the Japanese lamino-alveolo-palatal

fricative ([ɕ]) to the English post-alveolar [ʃ]. To sum up, more young females tend to pronounce [s]-like [ʃ] than the others.

As for the relationship between the presence of on-glide and the [s]-like [ʃ], two of the three participants in JG1 who pronounced quite [s]-like [ʃ] pronounced on-glide and the other 1 pronounced pure [s]. In JG1, 10 of the 24 participants and in JG2, two of the three participants who pronounced somewhat [s]-like [ʃ] at word-initial pronounced on-glide, whereas the other 14 in JG1 and one in JG2 pronounced pure [s]. Since these characteristics are quite evenly distributed, [s]-like [ʃ] and on-glide insertion seem to be two different phenomena. Since phonological contrasts tend to be kept clear (Ôno, 2003), I speculated that those whose [ʃ] is phonetically close to [s] would insert on-glide after [s] to maintain the clear distinction between /si/ and /ʃi/, but it turned out wrong. As for the relationship between [s]-like [ʃ] and [θ]-like [s], only one participant had both characteristics. On the same basis, I also speculated that both /s/ and /ʃ/ moved forward to become [θ]-like [s] and [s]-like [ʃ] respectively in order to maintain the clear distinction between these two phonemes. Or, if both [s]-like [ʃ] and [θ]-like [s] were weakened forms of /ʃ/ and /s/ respectively, people who weakened [ʃ] might also weaken (or front) [s] and vice versa. In terms of L2 acquisition, I conclude that those whose /ʃ/ which can be perceived as Eng/s/ do not necessarily pronounce /s/ which can be perceived as something between Eng/s/ and Eng/θ/. In this paper, I have not dealt much with the quality of [s] unless it has a role in a cross-linguistic issue, and whether my speculations are right or wrong is left to further studies.

## 3.4 Experiment 3: Listening

### *3.4.1 Objective*

The purpose of Experiment 3 is to determine whether Japanese speakers can hear the difference between Eng/si/ and Eng/ʃi/ in "non-English" contexts and whether Japanese speakers can distinguish between those syllables better or worse than other syllables. The participants listened to 11 nonsense words; six of them are monosyllabic and the other five are disyllabic.

### *3.4.2 Participants*

In Experiment 3, in addition to JG1 and JG2, another group (English Group or EG) was added. EG consisted of 20 native speakers of North American English who were not specializing in phonetics and phonology. One was from Hawaii living in Tôkyô in Japan and the others were Canadians living in British Columbia. Therefore, all 93 participants in JG1 except for one male in his 20's[32] (92 participants in total), 12 participants in JG2, and 20 participants in EG were involved in Experiment 3. Table 3.4 is the number of participants in each age group and gender in EG.

**Table 3.4**
*The number of participants in each group of age and gender in EG*

| Age | Male | Female | Total |
|---|---|---|---|
| 10-19 | 1 | 3 | 4 |
| 20-29 | 6 | 1 | 7 |
| 30-39 | 2 | 2 | 4 |
| 40-49 | | 2 | 2 |
| 50-59 | 1 | 1 | 2 |
| 60-69 | | 1 | 1 |
| Total | 10 | 10 | 20 |

*Note.* EG = native speakers of English

[32] As mentioned in footnote 28, two male participants in their 20's in JG1 had Down's syndrome. The results of one of them in Experiment 3 were discarded.

### *3.4.3 Verbal Stimuli*

[biː, niː, dʒiː, siː, tiː, ʃiː, niːbiː, miːʃiː, kiːtʃiː, dʒiːdiː, piːsiː]

The above stimuli were pronounced by a phonetically trained male native speaker of Canadian English in his 20's. My focus is on [siː] and [ʃiː], and the others are distracters made from English phonemes that also exist in Japanese. The important point of the Japanese writing system is that, as I showed in Table 2.2, some Chinese and other foreign originated syllables are written with two symbols in Japanese. Therefore, the foreign originated syllable [si] is usually written with two symbols "スィ" while the native syllable [ʃi] is written with one symbol "シ." In order to have the participants get used to this inconsistency, I mixed both syllables written with one symbol and those written with two symbols in the stimuli. Namely, [di] and [ti] are written with two symbols and the other distracters are written with one symbol. I randomized the first six monosyllabic stimuli and the other five disyllabic stimuli, and ended up with this order above. I showed the speaker those stimuli written with the English orthography in order to have him feel an English atmosphere: "bee," "nee," "jee," "see," "tee," "shee," "neebee," "meeshee," "keechee," "jeedee," "peesee." The speaker pronounced them with level tone in order to make them like one of the accent patterns in Japanese, namely *Hêbangata* or the non-accented-pattern, and not unnatural as an English intonation pattern as well. In addition, Japanese pre-pausal short vowels are typically followed by a glottal stop while pre-pausal long vowels typically fade out gradually (Vance, 2008). Since the way of Japanese pre-pausal short vowels is not natural in English, the speaker somewhat prolonged the vowel part in each stimulus in order to have the Japanese participants perceive it as a long vowel. For example, the stimulus [biː] is intended to be like Jap/biː/ with the non-accented-pattern as well as the

Eng/bi/. Recording was done in the UVic Phonetics Lab with Luna 1.1 inch large diaphram condenser microphone, M-Audio Firewire 410, with Praat set at 44100Hz. After the recording, another phonetically trained female native speaker of American English checked if those stimuli were properly pronounced, and she pointed out that the second stimulus [ni:] could be heard as [mi:]. In fact, over 40% of the participants in total, and over 50% in JG1 misheard it as [mi:]. Therefore, I discarded this stimulus and checked the other 10 stimuli. Since the last five stimuli were disyllabic, there were 15 syllable-initial consonants occurring. The sounds [ʃ], [s], [b], and [dʒ] occurred twice each, so there were 11 different consonants.

### *3.4.4 Procedure*

There were listening tasks, followed by a simple exit interview. In the listening part, the Japanese listeners wrote down what they heard with the Japanese orthography and the English-speaking listeners did it with the English orthography on the answer sheet. The Japanese listeners were not notified that the stimuli were pronounced by a Canadian person. It is because I tried to avoid other psychological factors that could affect their responses, such as fear of a foreign language and over-generalization. Technically what they did was to categorize each of those English consonants to one of the Japanese consonants.

On the other hand, the native English speakers were notified that the stimuli were basically nonsense English words. These English participants simply wrote down what English consonant they heard without thinking about the meanings.

All the participants were allowed to listen to the stimuli as many times as they wanted. This experiment was done with PowerPoint on a laptop, HP Pavilion dv2, and they listened to the stimuli with headphones, Pioneer SE-M390 (Frequency response:

5~29000Hz). After the listening tasks were done, I examined [siː], [ʃiː], and other distracters to determine the participants' overall listening ability and tendencies of errors. Their answers of vowels were ignored in this experiment.

In the exit interview, I asked the Japanese participants if they thought that the stimuli were pronounced by a Japanese person or a non-Japanese person. The main purpose of this question was to see how much the different qualities between Eng/ʃ/ and the Jap/ʃ/ ([ɕ]) affect the recognition of Japanese-speaking people in such short stimuli. After the question, they were notified that the stimuli were recorded by a Canadian.

What I had to be careful of here was that many scholars, such as Matsuzaki (1993) and Miyauchi (1993), pointed out that the way of writing loanwords with *Katakana* (Japanese syllabaries) is not clear, so that there can be a few variants for a single word. In fact, for example, one of the stimuli, /siː/, could be written in several different ways, such as "スィー" and "シィー" by the participants. Therefore, I encouraged them beforehand to write "スィー" to represent [siː] and "シー" to represent [ʃiː] in order to minimize the possibility that their mistakes were due to the variations of the Japanese writing system rather than due to their mishearing. Except for "スィー" and "シィー," the other sounds should be quite straightforward despite some variations of the symbols. For example, both "ビー" and "ビィー" stand for [biː] except the latter is not the standard way, or the participants who wrote the latter type intended to describe phonetic details, such as the transition part in the vowel or the like. The same is true for English orthography. For example, both "see" and "cee" stand for [siː].

I analysed whether there was any difference between native Japanese speakers' and native English speakers' performance, any correlation between the characteristics of their

perception and error patterns of the perception, what phonetic or phonological factors cause their errors, and so on. For the analyses, t-test and correlation test were used to examine relationships among variables.

*3.4.5 Results and Discussion*

Table 3.5 to 3.11 show the stimuli the participants heard and their responses. Table 3.5 to 3.8 are the responses for [ʃ] and [s]. Table 3.9 to 3.11 are for the other sounds in the distracters where [dʒi] and [bi] appeared twice and the others appeared only once.

**Table 3.5**
***The results of /s/ and /ʃ/ by 92 monolingual Japanese participants' responses (JG1)***

| Stimulus | **si** | **ʃi** | **isi** | **iʃi** |
|---|---|---|---|---|
| Response - % of correct responses (# of participants) | **s - 96.7% (N=89)** | **ʃ - 100% (N=92)** | **s - 77.2% (N=71)** | **ʃ - 98.9% (N=91)** |
| Response - % of responses (# of participants) | **ʃ - 3.3% (N=3)** | | **ʃ - 22.8% (N=21)** | **s - 1.1% (N=1)** |

**Table 3.6**
***The results of /s/ and /ʃ/ by 12 Japanese ESL learners' responses (JG2)***

| Stimulus | **si** | **ʃi** | **isi** | **iʃi** |
|---|---|---|---|---|
| Response - % of correct responses (# of participants) | **s - 100% (N=12)** | **ʃ - 100% (N=12)** | **s - 91.7% (N=11)** | **ʃ - 100% (N=12)** |
| Response - % of responses (# of participants) | | | **ʃ - 8.3% (N=1)** | |

**Table 3.7**
***The results of /s/ and /ʃ/ by 104 Japanese participants' responses (the total of JG1 and JG2)***

| Stimulus | **si** | **ʃi** | **isi** | **iʃi** |
|---|---|---|---|---|
| Response - % of correct responses (# of participants) | **s - 97.1% (N=101)** | **ʃ - 100% (N=104)** | **s - 78.8% (N=82)** | **ʃ - 99.0% (N=103)** |
| Response - % of responses (# of participants) | **ʃ - 2.9% (N=3)** | | **ʃ - 21.2% (N=22)** | **s - 1.0% (N=1)** |

Table 3.8
*The results of /s/ and /ʃ/ by 20 native North American English speakers' responses (EG)*

| Stimulus | si | ʃi | isi | iʃi |
|---|---|---|---|---|
| Response - % of correct responses (# of participants) | **s - 100% (N=20)** | ʃ - 85% (N=17) | **s - 100% (N=20)** | ʃ - 90% (N=18) |
| Response - % of responses (# of participants) | | s - 15% (N=3) | | **s** - 10% (N=2) |

Table 3.9
*The results of the other consonants by 92 monolingual Japanese participants' responses (JG1)*

| | | | | | Stimuli | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| dʒi1 | dʒi2 | ibi | bi | idi | pi | ni | mi | itʃi | ki | ti |
| **dʒ** | **dʒ** | **b** | **b** | **d** | **p** | **n** | **m** | **tʃ** | **k** | **t** |
| **100%** | 95.7% | 98.9% | 97.8% | 94.5% | 84.8% | 83.7% | 77.2% | 75.0% | 73.9% | 72.8% |
| **(92)** | (88) | (91) | (90) | (87) | (78) | (77) | (71) | (69) | (68) | (67) |
| | **tʃ** | ***v** | ***v** | **b** | **ɸ** | **m** | **n** | **ʃ** | **t** | **tʃ** |
| | 3.3% | 1.1% | 2.2% | 2.2% | 4.3% | 16.3% | 22.8% | 14.1% | 22.8% | 19.6% |
| | (3) | (1) | (2) | (2) | (4) | (15) | (21) | (13) | (21) | (180 |
| | **d** | | | **ɾ** | **t** | | | **s** | **tʃ** | **ts** |
| | 1.1% | | | 2.2% | 3.3% | | | 2.2% | 1.1% | 4.3% |
| | (1) | | | (2) | (3) | | | (2) | (1) | (4) |
| | | | | **dʒ** | **h** | | | **ts** | **b** | **ʃ** |
| | | | | 1.1% | 3.3% | | | 2.2% | 1.1% | 1.1% |
| | | | | (1) | (3) | | | (2) | (1) | (1) |
| | | | | | **b** | | | **dʒ** | NC | **z** |
| | | | | | 2.2% | | | 2.2% | 1.1% | 1.1% |
| | | | | | (2) | | | (2) | (1) | (1) |
| | | | | | **s** | | | ***b ʃ** | | **ʒ** |
| | | | | | 1.1% | | | 1.1% | | 1.1% |
| | | | | | (1) | | | (1) | | (1) |
| | | | | | **n** | | | ***t ʃ** | | |
| | | | | | 1.1% | | | 1.1% | | |
| | | | | | (1) | | | (1) | | |
| | | | | | | | | ***p ʃ** | | |
| | | | | | | | | 1.1% | | |
| | | | | | | | | (1) | | |

*Note*. * = phonotactically not allowed or not standard in Japanese; NC = no consonant

Table 3.10
*The results of the other consonants by 12 ESL learners' responses (JG2)*

| Stimuli | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| dʒi1 | dʒi2 | ibi | bi | idi | itʃi | ni | mi | ki | pi | ti |
| **dʒ** | **dʒ** | **b** | **b** | **d** | **tʃ** | **n** | **m** | **k** | **p** | **t** |
| **100%** | **100%** | **100%** | **100%** | **100%** | **100%** | **100%** | 91.7% | 91.7% | 83.3% | 83.3% |
| **(12)** | **(12)** | **(12)** | **(12)** | **(12)** | **(12)** | **(12)** | (11) | (11) | (10) | (10) |
| | | | | | | | **n** | **d** | **ɸ** | **tʃ** |
| | | | | | | | 8.3% | 8.3% | 8.3% | 8.3% |
| | | | | | | | (1) | (1) | (1) | (1) |
| | | | | | | | | | **h** | **ts** |
| | | | | | | | | | 8.3% | 8.3% |
| | | | | | | | | | (1) | (1) |

Table 3.11
*The results of the other consonants by 20 native North American English speakers' responses (EG)*

| Stimuli | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ibi | bi | idi | pi | ti | dʒi1 | dʒi2 | ki | ni | mi | itʃi |
| **b** | **b** | **d** | **p** | **t** | **dʒ** | **dʒ** | **k** | **n** | **m** | **tʃ** |
| **100%** | **100%** | **100%** | **100%** | **100%** | 95% | **100%** | 90% | 85% | 80% | 75% |
| **(20)** | **(20)** | **(20)** | **(20)** | **(20)** | (19) | **(20)** | (18) | (17) | (16) | (15) |
| | | | | | d | | **g** | **m** | **n** | **ʃ** |
| | | | | | 5% | | **5%** | 15% | 20% | 25% |
| | | | | | (1) | | (1) | (3) | (4) | (5) |
| | | | | | | | **t** | | | |
| | | | | | | | 5% | | | |
| | | | | | | | (1) | | | |

Figure 3.6 shows the overall percentage of the correct responses by monolingual Japanese speakers (JG1). If Japanese participants chose the Japanese counterpart consonant of the English stimulus, it was counted as a correct response.

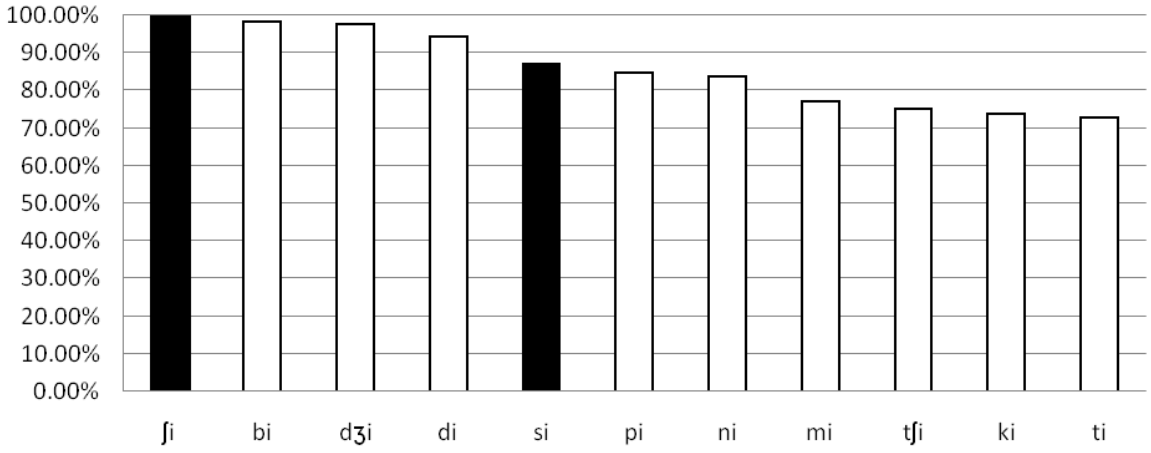


*Figure 3.6.* **Monolingual Japanese participants' responses (JG1)**

Table 3.12 shows the mean scores of all the participants' correct responses in each age and gender group. The stimuli regarding [s] or [ʃ] and those regarding the other consonants were counted separately. The maximum score regarding [s] and [ʃ] is 4 and that regarding other consonants is 11 because [s] and [ʃ] occurred 4 times in total and the other consonants occurred 11 times. Figure 3.7 shows the mean scores of only JG1.

**Table 3.12**
***Mean scores in each age and gender group***

| | JG1 92 monolingual JS | | JG2 12 ESL learners | | EG Native Speakers | |
|---|---|---|---|---|---|---|
| | Mean Score of s-ʃ Stimuli (out of 4) | Mean Score of Other Stimuli (out of 11) | Mean Score of s-ʃ Stimuli (out of 4) | Mean Score of Other Stimuli (out of 11) | Mean Score of s-ʃ Stimuli (out of 4) | Mean Score of Other Stimuli (out of 11) |
| 10-19 M | | | | | 4 | 11 |
| F | 4 | 10 | 4 | 10 | 4 | 10.33 |
| total | 4 | 10 | 4 | 10 | 4 | 10.5 |
| 20-29 M | 3.8 | 9.8 | 4 | 11 | 4 | 10.17 |
| F | 3.94 | 10.06 | 4 | 11 | 4 | 11 |
| total | 3.91 | 10 | 4 | 11 | 4 | 10.29 |
| 30-39 M | 4 | 9.5 | 3 | 10 | 3 | 10 |
| F | 3.83 | 10.16 | 4 | 10.8 | 3.5 | 10 |
| total | 3.9 | 9.9 | 3.83 | 10.67 | 3.25 | 10 |
| 40-49 M | 3.33 | 9.85 | | | | |
| F | 4 | 10.5 | 4 | 10 | 4 | 11 |
| total | 3.71 | 10.21 | 4 | 10 | 4 | 11 |
| 50-59 M | 3.5 | 9.25 | | | 4 | 10 |
| F | 3.8 | 10.3 | | | 4 | 10 |
| total | 3.71 | 10 | | | 4 | 10 |
| 60-69 M | 3.67 | 8.8 | | | | |
| F | 3.67 | 9 | | | 2 | 9 |
| total | 3.67 | 8.86 | | | 2 | 9 |
| 70-79 M | 3.2 | 7.8 | 4 | 9 | | |
| F | 4 | 8 | | | | |
| total | 3.33 | 7.83 | 4 | 9 | | |
| 80-89 M | 3 | 8.5 | | | | |
| F | 3 | 6 | | | | |
| total | 3 | 7.67 | | | | |
| overall M | 3.56 | 9.05 | 3.67 | 10 | 3.8 | 10.2 |
| F | 3.86 | 9.95 | 4 | 10.67 | 3.7 | 10.3 |
| overall mean | 3.73 | 9.54 | 3.92 | 10.5 | 3.75 | 10.25 |

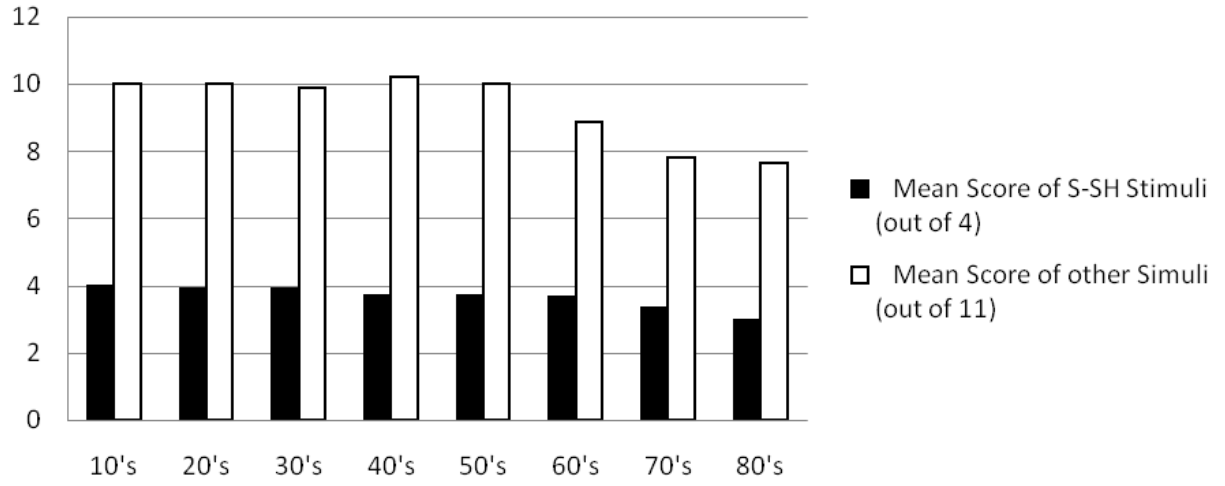


***Figure 3.7.*** **The mean scores of monolingual Japanese participants (JG1)**

Figure 3.8 is the scatter of the scores of both the stimuli regarding [s] and [ʃ] (white circles) and those regarding the other consonants (black circles) of each age group in JG1, JG2, and EG. For example, in JG1, the maximum score regarding the other consonants by participants in their 40's is 11, while the minimum score is 9. As well, the maximum score regarding the other consonants by participants in their 60's is 11, while the minimum score is 5. That means that the performance by participants in their 60's varies more than that by participants in their 40's. The scores of participants in their 70's vary from 6 to 10, and none in this age group scored 7.

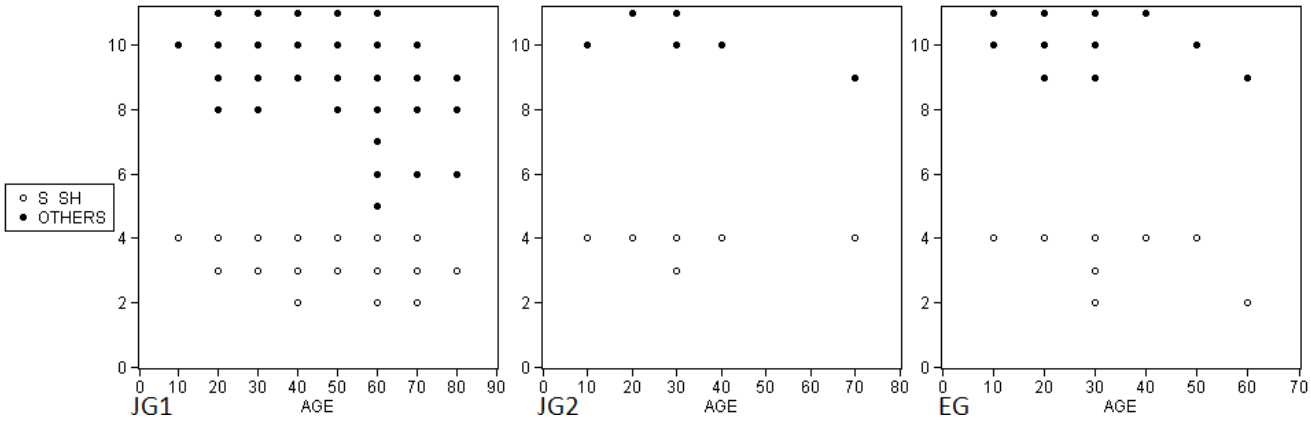


***Figure 3.8.*** **The scatter of the scores of both the /ʃ/ and /s/ stimuli and the other stimuli of each age group in JG1, JG2 and EG**

Table 3.5, 3.6, and 3.7 indicate that Japanese speakers are competent at recognizing [si:] and [ʃi:] in word-initial position. The most notable result is that all the Japanese participants in JG1 and JG2 correctly listened to [ʃi:] in word-initial position 100% of the time, and in word-medial position only one out of 92 monolingual participants confused [ʃi:] as [si:] but all the others got the right answer. In other words, Japanese participants correctly perceived these phonemes more than 99% of the time. More surprisingly, Japanese participants performed with [ʃi:] better than native English speakers did. These results contradict some previous studies mentioned in 2.7. Even more surprisingly, three out of 92 in JG1 and none in JG2, that is three out of 104 participants in total, confused [si:] with [ʃi:] in word-initial position. In other words, 97.1% of the Japanese participants correctly recognized [si:]. Even in JG1, still 96.7% of the participants correctly recognized word-initial [si:]. Among the 11 consonants, it is in 4th place after [ʃi:], [bi:] and [dʒi:]. This unexpectedly high percentage also completely contradicts the previous studies in 2.7. It seems to be generally believed that /s/ always appears as [ʃ] before [i] in Japanese, so that Japanese learners of English have difficulty in [s] before high front vowels, but these

results cannot be explained by the general belief. However, in word-medial position, the percentage of [si:] abruptly dropped to 77.2% among JG1 and 91.7% among JG2. Still, the overall percentages of [si:], as shown in Figure 3.6, is better than consonants with the features of nasals, voiceless affricates, and voiceless stops. This indicates that pronunciation problems with other phonemes than [ʃi:] and [si:] can be even more pronounced. Nonetheless, one cannot ignore that the 77.2% low score with the word medial phonemes and the notable difference between word-initial and word-medial positions. During the experiment, a few participants seemed not to be able to immediately decide whether [s] in the [pi:si:] stimulus should be categorized as Jap/s/ or Jap/ʃ/ and ended up choosing /ʃ/ after thinking. This implies that [s] in the [si:] stimulus is more clearly resembling Jap/s/ while [s] in the [pi:si:] stimulus is closer to the phoneme boundary for Japanese ears. As to why there is this difference, several explanations could be given.

The first possibility is whether [s] occurs word-initially or word-medially. This finding is the same as found in Masuda's (2003) experiment in which the participants also performed better in word-initial position.

The second possibility is the sound quality of [s] which is different between the [si:] stimulus and the [pi:si:] stimulus. Figure 3.9 (the [si:] stimulus) and Figure 3.10 (the [pi:si:] stimulus) show the difference.

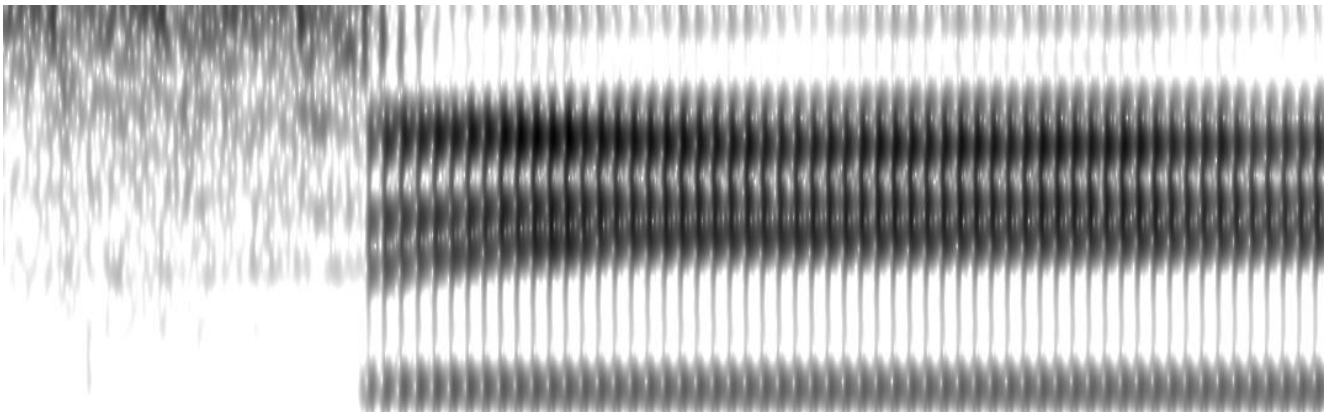


***Figure 3.9***. **The [si:] stimulus**
*Note*. 3 out of 104 Japanese participants recognized this as [ʃi:].

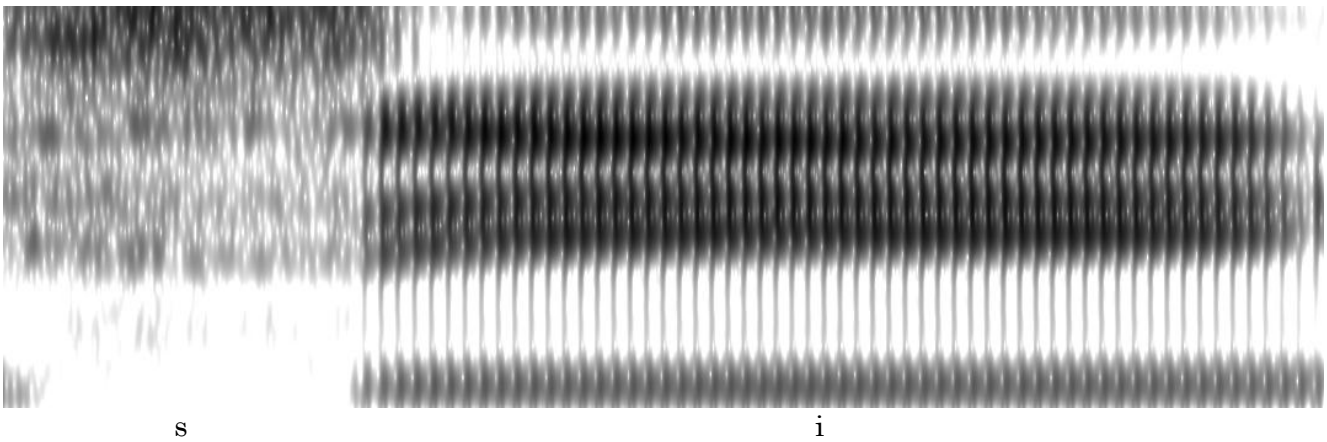


***Figure 3.10***. **The [si:] part in the [pi:si:] stimulus**
*Note*. 22 out of 104 Japanese participants recognized this as [ʃi:].

In Figure 3.10 the fricative part spreads out to the lower region while this does not happen in Figure 3.9. I selected 100msec from [s] in both stimuli. The centroid of it in the [si:] stimulus is 8929Hz, while that in the [pi:si:] stimulus is 7626Hz, calculated by Praat. As I mentioned in 2.6 with Figure 2.4, the centroid of Jap/ʃ/ (actually [ɕ]) is higher in frequency than that of Eng/ʃ/. As well, as mentioned in 3.3.8, centroid is one of the two factors that can separate Jap/ʃ/ and Jap/s/; the other factor is F2 onset of the following vowel. Therefore, when the fricative [s] comes a little lower, some Japanese participants categorized it as

Jap/ʃ/ rather than Jap/s/. Incidentally, in Experiments 1 and 2, the female Japanese judge in her 30's also sometimes found it difficult to decide if the participant's production should be categorized as /ʃ/ or /s/ when the other judge and I categorized it as perfect /s/. This indicates that the categorization of those 21 participants who answered /ʃ/ for the [piːsiː] stimulus is leaning more toward /ʃ/ than the categorization of those who answered /s/. Moreover, all three participants who categorized the [siː] stimulus (Figure 3.9) as /ʃiː/ also categorized the [siː] part in the [piːsiː] stimulus (Figure 3.10) as /ʃiː/. Figure 3.11 is a brief comparison of the continuum of the relevant speech sounds and their boundaries.

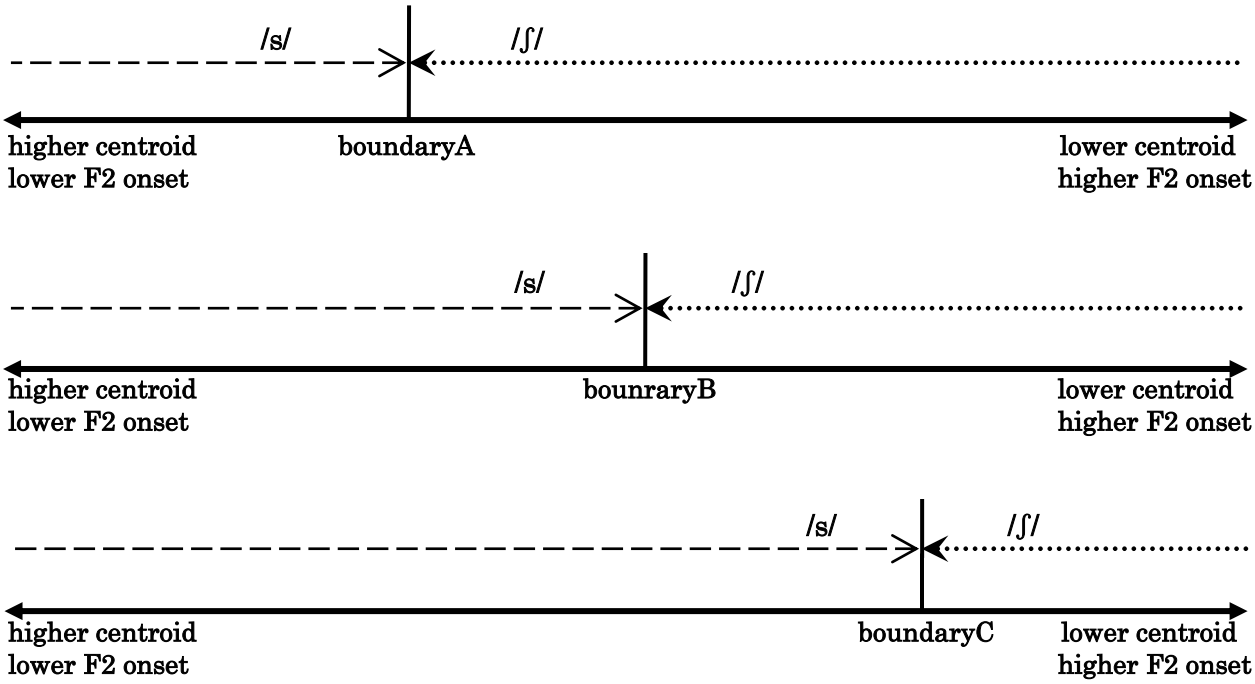


***Figure 3.11.*** **The continuum of the relevant speech sounds and their boundaries**

A good example of different phoneme boundaries is Voice Onset Time used in voiced and voiceless stops produced and perceived by French and English-speaking people. Although both English and French have /t/ and /d/, the sound that may be categorized as /d/ by English ears may be categorized as /t/ by French ears because of different voice onset time

(VOT) (Grenon, 2008). Likewise, phoneme boundaries between /s/ and /ʃ/ must be different between English and Japanese in centroid and F2 onset of the following vowel. Of course there are individual variants among speakers of the same language.

In addition to centroid, transition between the two [si:] syllables could also have affected the perception. In English, the so-called monophthongs display considerable vowel-intrinsic spectral, whereas in Japanese both short and long vowels are produced as pure monophthongs (Akahane-Yamada, et al. 2001). Japanese speakers are so used to pure monophthongs that they could be sensitive to vowel-intrinsic spectral change. In fact, a comparison of F2 in both Figure 3.9 and 3.10 shows that F2 in Figure 3.9 curves more. Vowel-intrinsic spectral change could affect the participants' perception. I also looked at each value of F2s and F3s on the 1st, 5th, 10th, and 15th pulse in both [si:] syllables by using Formant Listing on Praat. The measured point on each pulse was the nearest 0 crossing from the highest spike. According to Akiyama (2009), in Japanese, the most preferable speech speed is eight moras per second; in other words, one mora in about 0.125 second. The duration of those 15 pulses in both [si:] syllables is nearly 0.125 second, which should be long enough to be recognized. Table 3.13 shows each value of F2 and F3.

**Table 3.13**
***Each value of F2 and F3 in the [si:] stimulus and the [si:] part in the [pi:si:] stimulus***

| | | 1st | 5th | 10th | 15th |
|---|---|---|---|---|---|
| [si:] | F2 | 1909(Hz) | 2047 | 2161 | 2249 |
| | F3 | 2392 | 2619 | 2700 | 2761 |
| [pi:si:] | F2 | 2052 | 2217 | 2227 | 2256 |
| | F3 | 2494 | 2765 | 2829 | 2828 |

The higher F2 onset in the [pi:si:] stimulus could have affected the participants' perception. When [s] is sandwiched between [i]'s, it is likely to be articulated closer to the palate influenced by both [i]'s. As well, in the [si:] stimulus, both F2 and F3 kept going up from

the 1st to 15th pulses, while in the [pi:si:] stimulus, they rise abruptly only from the 1st pulse to the 5th pulse, but after that they are quite stable. Therefore, the transition in the [si:] stimulus could be detectable for Japanese ears. As well, both F2 and F3 in the [si:] stimulus start from over 100Hz lower than those in the [pi:si:] stimulus and end up at the more resembling values of frequency. For the participants who pronounced [sɯi] or [swi], the presence of the glide could be one of the cues for the distinction between /si/ and /ʃi/. They could have recognized the transition in the [si:] stimulus as a glide and categorized the stimulus as /si:/, while the transition in the [pi:si:] stimulus was not long enough to be recognized as a glide and they leaned more toward /ʃi:/. In fact, among the 22 participants who categorized the [si:] in the [pi:si:] stimulus as /ʃi:/, 16 (72.7%) of them pronounced on-glide and six (27.3%) of them always pronounced [si]. On the other hand, among the 82 participants who categorized the [si:] in the [pi:si:] stimulus as /si:/, 35 (42.7%) of them pronounced [sɯi] or [swi] while 47 (57.3%) of them always pronounced [si]; the ratio is reversed. It seems those who pronounced on-glide were likely to hear /ʃi:/ when there was a less detectable transition. In short, the audibility of the transition can be one of the cues for distinguishing between /ʃi/ and /si/ for some Japanese speakers. The combination of lower centroid and a more audible transition, or higher F2 onset, can be considered to make the participants hear Jap/ʃ/.

**Table 3.14**
***The relationship between the probability that [si:] with small transition is categorized as /ʃi:/ or /si/ and the probalility of occurence of on-glide***

| Perception | Production |
|---|---|
| **22** participants who categorized [si:] in the [pi:si:] stimulus as **/ʃi:/** | **[sɯi] or [swi]** – 16 (**72.7%**) |
| | [si] – 6 (27.3%) |
| **82** participants who categorized [si:] in the [pi:si:] stimulus as **/si:/** | [sɯi] or [swi] – 35 (42.7%) |
| | **[si]** – 47 (**57.3%**) |

The third possibility, in addition to the positional difference and the phonetic sound quality, could simply be physiological hearing differences. Older participants may have worse hearing than the younger participants. In Figure 3.7 and 3.8, the scores of both [si:] and [ʃi:] stimuli and the other consonant stimuli go down after the age of 60. I could also speculate that people who made errors in the stimuli regarding [si:] and [ʃi:] were also likely to make errors in the other consonant stimuli because their physiological hearing was not as good as those who did not make errors regardless of their age. The relationship between physiological hearing and L2 sound recognition may be a future study topic.

The fourth possibility could be that some people tended to pay less attention to details, less interested in pronunciation, are less like auditory learners, and so on. In other words, if those people made errors in both stimuli, there may be a difference from those who made fewer errors in sound recognition skills. Of course, that does not mean that those who made more errors are not good language learners. Table 3.15 shows the comparison between the mean scores of the other consonant stimuli by 23 participants who incorrectly categorized either stimuli regarding [si:] and [ʃi:] at least once, and the 81 participants who got correct responses for all the four [si:]-[ʃi:] stimuli.

**Table 3.15**
***The mean scores of the other consonant stimuli (the maximum score is 11)***

| | Mean |
|---|---|
| **23** participants who incorrectly categorized the [si:]-[ʃi:] stimuli at least once | 9.39 |
| **81** participants who correctly categorized all the [si:]-[ʃi:] stimuli | 9.73 |

Those who categorized the [si:]-[ʃi:] stimuli correctly seem to perform slightly better with the other stimuli; the difference of the mean values is 0.34. However, the difference of the scores is not significant based on the analysis with two sample t-test with unequal

variances ($p > 0.1$). As well, the correlation between the scores of the [siː]-[ʃiː] stimuli (the maximum is four) and those of the other stimuli (the maximum is 11) is not significant, either ($p > 0.1$). Therefore, this hypothesis does not seem plausible; those who have trouble with [siː] and [ʃiː] do not necessarily have trouble with other consonants, while those who have trouble with other consonants do not necessarily have trouble with [siː] and [ʃiː]. In fact, six of the 22 participants who incorrectly answered /ʃiː/ for the [piːsiː] stimulus answered all the other distracters correctly, and another six of the 22 made only 1 error with the distracters. More than half of them did well with the distracters. The difference of their categorization regarding [siː] and [ʃiː] is probably from different causes than those of the other consonant stimuli. The frequency of the energy and transition seem to be a more trustworthy cue than their overall sound recognition skill. Again, note that the words "correct" and "incorrect" here are used in the criteria of whether their responses and the Japanese counterparts of the stimuli matched or not.

Finally, there could be some non-phonetic factors. For example, although I urged them to write "スィ" for [si], there might still be some possibility of the influence by the inconsistency of the symbol-sound correspondence in Japanese. In other words, a few of them actually might have perceived correctly but wrote wrongly. As well, one of our judges pointed out that some participants might have associated the [piːsiː] stimulus with the Japanese words /piːʃiː/ (PC, computer) and were inclined to answer /ʃiː/.

To sum up, I can assume that the factors below made the Japanese participants categorize differently. The combination of centroid and audibility of the transition of F2 onset is especially considered to be influential.

- Word-initial or word-medial position

- Frequency of the centroid in the fricative
- Audibility of transition or F2 onset of the following vowel

### *3.4.6 Comparison with Native English Speakers*

What is intriguing is that native English speakers also confuse /s/ and /ʃ/ as it is shown in Table 3.8. Therefore, the belief that the /s/ and /ʃ/ confusion is a major weakness of Japanese ESL/EFL learners is not empirically substantiated. However, there is a noticeable difference between the Japanese and English speakers' error patterns. While the Japanese participants almost always categorized [ʃ] as /ʃ/ but occasionally categorized [s] as /ʃ/, the English-speaking participants always categorized [s] as /s/ but occasionally categorized [ʃ] as /s/. In other words, the native speakers of Japanese are biased toward /ʃ/; in contrast, the native speakers of English are biased toward /s/. As mentioned in 2.6, Beckman et al. (2009) stated that adult English listeners accept a wider range of centroid for /s/ than adult Japanese listeners. Therefore, this cross-linguistic difference of the phoneme boundary between /ʃ/ and /s/ is quite understandable. In Figure 3.11, I could roughly pinpoint that native speakers of English have boundary C, native speakers of Japanese generally have boundary B, and the Japanese participants who categorized [s] as /ʃ/ in this experiment have boundary A. Only one Japanese monolingual female participant in her 50's in JG1 categorized one of the [ʃi:] stimulus (word-medial) as /si:/ and always categorized [s] as /s/. In the production tasks, she produced pure [si] with no glide, as well. Perhaps her boundary is closer to the English speakers' although she had not lived in an English-speaking country.

One more important point is that the Japanese participants had the most trouble with [ti:] as it is shown in Table 3.9, 3.10, and Figure 3.6, while the English-speaking

participants perfectly recognized it as it is shown in Table 3.11. The boundary between voiced stops and voiceless stops is decided by VOT as well as tenseness of the lips or tongue; in voiceless stops, muscles get tenser. In Japanese, word-initial voiceless stops typically have VOTs in between those for prototypical voiceless unaspirated stops and prototypical voiceless aspirated stops in other languages (Vance, 2008). In Japanese consonants, muscles are generally less tense than in English consonants (Inozuka, et al., 2009). Therefore, the phonological boundary between voiced and voiceless stops should be different between English and Japanese. In fact, the Japanese participants had much more trouble with English voiceless stops than their voiced counterparts.[33] According to the results of this study, acquiring the boundary between voiced and voiceless stops in English seems to be more problematic for Japanese speakers than acquiring the boundary between /s/ and /ʃ/ in English. Moreover, in terms of place of articulation, Jap/t/ is lamino-alveolar, which means that the front of the tongue comes in contact with the alveolar ridge while Eng/t/ is apico-alveolar (Koizumi, 2003; Vance, 2008). The difference in place of articulation makes Eng/t/ difficult to recognize for Japanese speakers. Interestingly, the voiced counterpart [di:] stimulus was easily recognized by the Japanese participants. Since Eng/t/ has longer VOT and more tenseness, these characteristics may make the difference of the place more notable and the Japanese participants did not perceive Eng/t/ as Jap/t/. The point here is that the more infamous Eng/si/ is actually less problematic than the less infamous Eng/ti/.

Both Japanese and English participants' performance on [m], [n], and [tʃ] is relatively poor; JG1 scored 77.2% for [m], 83.7% for [n], and 75% for [tʃ], whereas EG scored 80% for [m], 85% for [n], and 75% for [tʃ], as shown in Table 3.9 and 3.11. The fact that

---

[33] Since my focus was on [s] and [ʃ], stops were made just as distracters. I did not make any stimuli with [g].

speakers of two languages have the same characteristics probably cannot be considered as a cross-linguistic issue, and I will leave it for future studies.

### *3.4.7 Exit Interview*

After the listening tasks, the Japanese participants were orally asked whether they thought the stimuli were pronounced by a Japanese person or a non-Japanese person. The main purpose is to find out whether the quality of Eng/ʃ/ sounds foreign to Japanese ears. If it does, acquisition of Eng/ʃ/ should be considered relatively important. 67 in JG1 and eight in JG2 thought that the speaker was not Japanese, while 25 in JG1 and four in JG2 thought that he was Japanese. Some of them could not really judge it, but they commented that if they had only two possibilities to choose from, they would answer "non- Japanese." Despite all my arrangements for making the English verbal stimuli fit in the Japanese phonology as much as possible, 72.1% of the Japanese participants suspected that the stimuli were recorded not by a Japanese speaker. Since the speaker is a phonetician, he pronounced the stimuli very clearly so that the characteristics of English may have stood out. The participants were not required to tell the reason, but 38 participants in JG1 and six in JG2 commented why they thought the speaker was not a Japanese person. The others did not comment, or said that they could not explain why. Two in JG1 and one in JG2 gave two different comments, so there are 40 comments from JG1 and six comments from JG2. Their comments are organized in Table 3.16.

**Table 3.16**
***Reasons why they thought that the stimuli were recorded by a non-Japanese person***

| N of participants JG1 | N of participants JG2 | Reasons why they thought that the stimuli were recorded by a non-Japanese person. |
|---|---|---|
| 7 | 2 | His voice quality is not like a Japanese person's. His voice is thicker, lower, or less sharp. |
| 6 | 1 | The vowel quality is different; there is a small vowel. For example, the stimuli [bi:], [pi:], and [ʃi:] sound like [bɯi:], [pɯi:], and [ʃii:] respectively, and so on. |
| 5 | 2 | It is somewhat difficult to exactly transcribe with the Japanese orthography or to exactly categorize as Japanese sounds. |
| 3 | 0 | The consonants are different. The consonants are clear and easy to hear. The consonants are long; for example, the stimuli [mi:] and [ni:] sound like [mmi:] and [nni:] respectively.[34] |
| 3 | 0 | The intonation is not like Japanese. Perhaps in Japanese the ending usually falls down. |
| 3 | 0 | Each sound is clear or beautiful. |
| 2 | 1 | [si:] and [ti:] sound different from those in Japanese. (They did not explain how different.) |
| 2 | 0 | [si:] sounds foreign. He is good with [si:], like a CD of English pronunciation material. |
| 2 | 0 | The 1-syllable stimuli sound like Japanese, but the 2-syllable stimuli sound foreign. |
| 1 | 0 | He is good with *Yôon*.[35] |
| 1 | 0 | I used to record voice as my job, so I can tell the difference between Japanese speakers and non-Japanese speakers. |
| 1 | 0 | His voice goes strong and weak. |
| 1 | 0 | Japanese should sound stickier. |
| 1 | 0 | His pronunciation is nasal. |

[34] Since length of consonants is phonemic in Japanese, Japanese speakers may be sensitive to it.
[35] See footnote 9.

| | | |
|---|---|---|
| 1 | 0 | I associated the stimulus [ni:] with the English word "knee." |
| 1 | 0 | It sounds like he is a foreign person who is proficient at Japanese. |

The most commonly reported reason was that his voice quality was not like a Japanese person. Wilson (2006) pointed out that each language has its "underlying or default posture" or "articulatory setting" (p.1). One of the innovative English pronunciation teaching materials in Japan written by Uekawa (2007) mentioned that English sounds come more from the throat than Japanese sounds. In linguistic terms, he mentioned lowered larynx voice; he pointed out language-specific laryngeal settings. People in general (especially non-musical people) tend to express "low voice" or "high voice" based on formants, rather than fundamental frequency (Teshigawara, 2010). Some of the participants expressed that the man who pronounced the stimuli had a low voice, which actually must have meant that his default laryngeal position is perhaps lower than that of typical Japanese speakers so that the formants are lower. The second most commonly reported reason was that his vowels sounded like two sounds. What they call "a small vowel" probably means "transition" in a linguistic term. This can be interpreted that Japanese ears are so used to pure monophthongs that they are quite sensitive to the quality change within a vowel. Two in JG1 and one in JG2 named specific stimuli, [si:] and [ti:], and two in JG1 named [si:]. As well, one in JG1 pointed out "*Yôon*" which are namely [si:], [ti:], and [di:] in the stimuli. They did not tell in what way the stimuli sounded foreign. Perhaps that is phonetic differences, such as aspiration in [t], no clear glide insertion in [si:], transition in the vowel part, and so on. Or perhaps these syllables appear only in foreign originated words in Japanese and the participants simply associated them with a foreign person. Contrary to my prediction, none of them pointed out the quality of [ʃ]. That

indicates that in those short stimuli, the English-like vowel quality and the English-like laryngeal setting seem much more salient than the difference between Eng/ʃ/ and Jap/ʃ/ for Japanese ears, at least at the conscious level. Therefore, I argue that although pronunciation textbooks in Japan often mention the sound quality of Eng/ʃ/ and Jap/ʃ/, such a mention might be beyond beginner learners' understanding if this difference is perceptually not salient enough to Japanese ears.

Incidentally, one participant in JG2 guessed that the speaker was a Japanese who pretended to be a foreigner. Some of the other participants who thought that the stimuli were recorded by a Japanese person commented that those stimuli were clear and easy to hear.

### *3.4.8 Other Comparisons*

According to Table 3.12, in comparison between JG1 and JG2, JG2 performed better. Since no participants in JG2 were in their 50's, 60's, and 80's, I compared the participants in JG1 in their teens, 20's, 30's, 40's and 70's with those in JG2 by using two sample t-test with unequal variances. Interestingly, JG2 performed significantly better with the stimuli with other consonants than [s] and [ʃ] than JG1 did ($p < 0.01$). It can be because JG2 were more used to the English phoneme boundaries and were less confused. On the other hand, the overall scores of the [s] and [ʃ] stimuli were not significantly different between JG1 and JG2 ($p > 0.1$). However, in word-medial position, only 77.2% of the participants in JG1 correctly perceived [si:] while 91.7% of the participants in JG2 correctly perceived it. It may be because JG2 were more used to the English phoneme boundaries between /s/ and /ʃ/.

Another comparison is the gender difference. According to Table 3.12, females usually performed better than males within the same age group. However, in JG1, when

the differences between means for both the stimuli regarding [s] and [ʃ] and those with other consonants within each age group were analyzed by using two sample t-test with unequal variances, there were no significant differences found[36] ($p > 0.1$ in all the age groups). In JG2 and EG, I combined all the age groups because some age groups did not have sufficient participants and used t-test. The gender differences in both JG2 and EG were not significant ($p > 0.1$ in JG2 and EG).

### 3.5. Conclusion

Despite the general belief that Japanese ESL learners cannot distinguish between [s] and [ʃ] before high front vowels, in both production and perception, Japanese speakers clearly recognize [si] and [ʃi] as two different entities regardless of their age and gender. Although the sample size 93 + 12 may not be big enough to generalize, it is worthy of serious attention that all the participants have distinguished [si] and [ʃi] with no exceptions so far. As well, the fact that there are some scholars who accept /si/ in the Japanese phonological system cannot be ignored. I could say that the results are valid. More surprisingly, [ʃi] produced by an English native speaker was perceived more accurately by both monolingual Japanese speakers and Japanese ESL learners than by native English speakers. When the Japanese participants saw the written material, they were less confused by the tongue-twister-like stimulus than when they heard the sound material. On the other hand, although both English-speaking people and Japanese-speaking people recognize these sounds as different entities, the phoneme boundary is different between English and Japanese. Specifically, English-speaking people categorize

---

[36] Teens and 20's were combined, so were 70's and 80's because there was only one female in teens, 70's and 80's.

wider range of sound as /s/ than Japanese-speaking people do. Especially, word-initial [ʃ] produced by some Japanese speakers could be categorized as Eng/s/ or somewhere between Eng/s/ and Eng/ʃ/ by English ears. /ʃ/ produced by a few Japanese participants, especially young females, was quite close to [s] while still retaining the contrast from /s/. Whether this characteristic is an issue within Japanese or that regarding ESL/EFL is a question for the future. As for perception, Eng/s/ in the #_i context was perceived as Jap/s/ 97.1% of the time and as Jap/ʃ/ 2.9% of the time by the Japanese participants, while Eng/s/ in the i_i context was perceived as Jap/s/ 78.8% of the time and as Jap/ʃ/ 21.2% of the time. This asymmetry is probably due to different centroid of the fricative and audibility of transition or F2 onset of the following vowel. However, this perceptual difference in comparison with English-speaking people is still less problematic than the English voiceless stops, especially /t/. Another difference from English is that about 50% of the Japanese participants sometimes or always add on-glide in the [si] syllable, like [sɯi] or [swi], while they pronounce pure [ʃi] with no glide inserted. Age or gender and this tendency seem to have a weak correlation; older people are slightly more likely to add on-glide, and males are also slightly more likely to add on-glide. When a new mora is fixed in the Japanese phonological system, the process from two moras to one mora is often seen. For example, /ɸi/ started from [ɸɯi] and eventually became [ɸi]. Since [si] has not been fully accepted in the Japanese phonological inventory by many linguists, those 50% of the Japanese participants are somewhere between [sɯi] and [si] while the others have already completed the process and pronounce pure [si]. Since a few of them very strongly pronounced the glide, it needs to be examined whether there might be a possibility that they find it difficult to distinguish between Eng/s/ and Eng/sw/ before high front vowels, such as “seat” vs. “sweet” and “sift” vs.

"swift," rather than "seat" vs. "sheet" and "sift" vs. "shift." For further studies, it is needed to examine how native Japanese-speaking people and native English-speaking people perceive some characteristic sounds produced by Japanese speakers, such as [si]-like /ʃi/, [swi]-like /si/, and so on.

# Chapter Four

# EXPERIMENT B: [si] AND [ʃi] IN ENGLISH CONTEXTS

## 4.1 Research Question

The study in Chapter Three shows that Japanese speakers can distinguish between [s] and [ʃ] before the high front vowel [i] in the Japanese contexts in both production and perception. Recall that in Fukushima, et al.'s (2008), over 80% of Japanese learners of English cannot correctly distinguish between those two sounds in the English contexts. If Japanese speakers can produce and perceive the difference between [si] and [ʃi], what then would be the sources of Japanese speakers' challenges in mastering the distinction in their English production?

To support the issue of [si] and [ʃi], I will also examine whether Japanese ESL learners can pronounce the English diphthongs /ei/ (Eng/ei/) and /ou/ (Eng/ou/) as in "able" and "hope" respectively. Thayne (2009) pointed out Eng/ei/ and Eng/ou/ as problematic vowels for Japanese ESL learners. On the other hand, according to Vance (2008), /ei/ and /oɯ/ exist in Japanese. Jap/ei/ and Jap/oɯ/ can be the counterparts of Eng/ei/ and Eng/ou/. If Japanese speakers produce and perceive /ei/ and /oɯ/ in the Japanese contexts, what then would be the sources of Japanese speakers' challenges in mastering Eng/ei/ and Eng/ou/ in their English production? Is the issue regarding /ei/ and /ou/ triggered by the same sources as the issue of [si] and [ʃi]?

## 4.2 Prediction

It is well known that a second language learner tends to transfer L1 constraints to L2 and may end up with the L1 constraint in his or her inter-language (IL) (Gass & Selinker, 2009). Since IL varies from learner to learner, the constraint ranking may also

vary. Figure 4.1 provides an example of possible L1 interference errors by a Japanese learner of English who pronounces "read" as [ɹiːdɯ]. I use the tableau of the theoretical instrument from Optimality Theory, which can be illustrated as follows; the candidate (c) is optimal.

| | | L1 Phonological Constraints | |
|---|---|---|---|
| **Target in L2**: /ɹiːd/ "read" | Ident-IO (ɹ) | No Coda (Obstruent) | *ɹ |
| a. ɹiːd | | *! | * |
| b. ɾiːdɯ | *! | | |
| c. ☞ ɹiːdɯ | | | * |

***Figure 4.1.*** **An example of the process of a Japanese ESL learner's pronunciation error in IL**
*Note.* Ident-IO = All the segments in input are identical to those in output; No Coda (Obstruent) = Obstruents, such as [d] and [k], cannot occur in coda position; *ɹ = [ɹ] is not allowed; * = the candidate violates the constraint (When there are two symbols of *, the constraint is violated twice.); ! = the candidate is not optimal; ☞ = the candidate is optimal.

However, a learner may make an error in the way of Figure 4.1 only when the learner knows exactly what the target is and when he or she attempts to pronounce the target. Some L2 learners may not know the proper target sounds. Based on Fraser (in press) and Nogita (2010), I hypothesize that Japanese ESL learners who confuse [s] and [ʃ] do not know when to pronounce [s] and [ʃ]. From my experience as an EFL learner, aforementioned *Katakana*-English pronunciation is predominantly used in Japan, and the process of Japanese ESL learners' errors may go through the processes shown in Figure 4.2.

| | Stereotypical L2 Flavours | |
|---|---|---|
| ***Katakana*-English**: /ɾi:do/ "read" | [do] sounds Japanese. [dɯ] is more exotic.[37] | The more rhotacized, the more English-like. |
| a. ɹi:do | *! | *** |
| b. ɾi:dɯ | | ****! |
| c. ☞ ɹi:dɯ | | *** |

***Figure 4.2*. A process of a Japanese ESL learner's pronunciation error in IL based on my hypothesis**
*Note*. [do] sounds Japanese. [dɯ] is more exotic. = [do] must be changed to [dɯ]; The more rhotacized, the more English-like. = All the segments in output must be [ɹ]; * = the candidate violates the constraint (When there are two symbols of *, the constraint is violated twice and so on.); ! = the candidate is not optimal; ☞ = the candidate is optimal.

Between Figure 4.1 and 4.2, although the outputs happen to be the same, the processes are completely different. Perhaps "Stereotypical L2 flavours" can also be considered as personal preferences of certain language-specific phonetic characteristics. Therefore, the major difference between Figure 4.1 and 4.2 is that Figure 4.2 starts from L1 and does not go through the process of L2 at all. As mentioned in footnote 25, *Katakana*-English pronunciation has often gone through complicated loanword adaptation processes. In addition, since the constraints in Figure 4.2 are not phonological, but rather may be psychological factors, the outputs may not even consist of discrete phonological sounds as linguistic sounds should be categorized as phonemes (D. McKercher, personal communication, 2009). Therefore, learners who rely on the process of Figure 4.2 may not know the proper target sounds, nor may they have phonological awareness in L2. If I assume that the learner's errors are based on the process in Figure 4.1 when it is actually based on Figure 4.2, I will end up with a wrong conclusion. This experiment aims to figure out whether Japanese ESL learners' mispronunciations are caused by not knowing when to pronounce [s] and [ʃ].

---

[37] The syllable [do] appears in native Japanese words, while [dɯ] is often considered as a foreignism in Japanese (Vance, 2008).

## 4.3 Procedure

Twenty-seven participants had a pre-test, a 40-munite lesson, an interview, and a post-test. All of these tasks took about 1.5 hours in total and were done on the same day. These 27 participants did not gather on the same day; instead, they were divided into eight groups, each of which consisted of a few participants, but they all did the same tasks. In the pre-test, the participants were recorded reading an English passage with no pronunciation tips (See Appendix A: Pre-test). They were told that the focus of this test was on each word, so that they did not have to read fluently. After the reading, the participants were taught the basic English phonological system and some basic symbol-sound correspondence rules such as, "a"-/ei/, "s"-/s/, and "sh"-/ʃ/ (See 4.6). The lesson lasted 40 minutes during which the participants were also interviewed to find out their awareness of the English phonological system and the symbol-sound correspondence (See 4.7). No articulation explanations were given during the lesson. After the lesson, they did a post-test by reading the same passage, but this time, the sounds at issue were underlined with the pronunciation indicated alongside with the English orthography for consonants and with the Japanese orthography for vowels (See Appendix B: Post-test).

Recording was done in the UVic Phonetics Lab with Luna 1.1 inch large diaphram condenser microphone, M-Audio Firewire 410, with PRAAT set at 44100Hz. Each participant was recorded in a small booth, so that the rest of the participants could not hear it. By doing so, each reader felt comfortable and was not influenced by other participants' pronunciation.

In order to examine the differences between the pre-test and the post-test, two-tailed paired t-test was used.

## 4.4 Participants

The 27 participants were all Japanese ESL learners. I assigned them the labels P1 (for Participant 1) through to P27 (for Participant 27). The participants were from two different ESL schools in British Columbia, except P13 was a first year undergraduate student. Table 4.1 shows the summary of the participants.

**Table 4.1**
*Participants in Experiment B*

| | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Gender | F | F | F | F | F | F | F | F | F | F |
| Age | 20's | 10's | 20's | 10's | 10's | 10's | 20's | 20's | 20's | 20's |
| LOR (month) | 2 | 2 | 2 | 2.5 | 2 | 2 | 2 | 2.5 | 2.5 | 2.5 |
| Level | 1 | 2 | 3 | 4 | 1 | 3 | 1 | 2 | 3 | 2 |

| | P11 | P12 | P13 | P14 | P15 | P16 | P17 | P18 | P19 | P20 |
|---|---|---|---|---|---|---|---|---|---|---|
| Gender | F | F | F | M | M | F | F | F | F | F |
| Age | 10's | 10's | 20's | 20's | 20's | 30's | 20's | 30's | 20's | 20's |
| LOR (month) | 3 | 3 | 18 | 2 | 1 | 1 week | 1 | 8 | 5 | 5 |
| Level | 3 | 3 | 7 | 3 | 4 | 2 | 2 | 2 | 2 | 3 |

| | P21 | P22 | P23 | P24 | P25 | P26 | P27 |
|---|---|---|---|---|---|---|---|
| Gender | F | M | F | F | F | M | F |
| Age | 20's | 30's | 20's | 20's | 30's | 20's | 20's |
| LOR (month) | 4 | 5 | 3 | 9 | 8.5 | 8 | 2 |
| Level | 2 | 3 | 3 | 4 | 4 | 5 | 4 |

*Note*. This table shows each participant's gender, age, length of residence (LOR) in an English-speaking country, and the level in the ESL school; 1 is the lowest class, 6 is the highest class and 7 is a university undergraduate program.

P13 had lived in New Zealand for 11 months and she arrived in Canada seven months prior to this experiment. P18 lived in the Great Britain for half a year before coming to Canada and arrived in Canada two months prior to this experiment. P16 arrived in Canada one week prior to the experiment. All of the participants had taken English classes for six years in junior and senior high school in Japan. A few of them were from the same universities or the same company but all of them had graduated from different high schools. In other words, all of them had taken the six-year English classes in different high schools.

### 4.5 Rating

A phonetically trained native speaker of Canadian English and I judged the recordings. The judges checked if the [s] and [ʃ] tokens produced by the participants before the high front vowels, [i] and [ɪ], can be categorized as the English /s/ and /ʃ/ respectively. The judges also observed /z/ preceding high front vowels, /ei/ and /ou/ as references since these sounds are likely to be affected by *Katakana*-English pronunciation. The criterion of the judgment was whether the participants hit the target sounds at the phonemic level. Their productions which were phonetically not the typical target sound but still fit in the range of the target phoneme were considered correct. When their productions were off the target, what they actually pronounced was transcribed. When the participants corrected themselves during recording, the judges picked up their corrected sound no matter if their first choice was actually correct. For example, if one pronounced "see" as [si:] first and immediately re-pronounced it as [ʃi:], the judges considered his or her production as [ʃi:]. As well, if one pronounced "see" as something like [ʃsi:], the judges considered his or her final decision as [s] because the participant changed his or her mind while pronouncing the word. When a sound was perceptually difficult to transcribe with one symbol, the judges wrote two IPA symbols. For example, when one's pronunciation was something between [ʃ] and [s], the judges wrote "s-ʃ" and counted it as an error. The judges assumed that the participants' target language was Canadian English, or North American English.

### 4.6 Lesson

I explained the following points with as few technical terms as possible, and did not tell the participants anything about articulation. The lesson was done in Japanese, the participants' first language.

- Just like Japanese has its symbol-sound correspondence as represented by the 50-Sound-Chart,[38] English does have one as often represented by "phonics" or the like. The fact that the English orthography is called "deep orthography" (Fayol & Pacton, 2004) does not mean there is no relationship between spelling and sounds. Look at the spellings much more carefully and never refer to *Katakana*-English. Moreover, the English spelling rule is almost completely different from the rule of the Japanese *Rômaji* (Romanization of Japanese). In addition, letter knowledge precedes phoneme segmentation (Carroll, 2004).
- Basically, the letter "s" says /s/, "sh" says /ʃ/, "a" says either /ei/ or /æ/, and "o" says either /ou/ or /ɑ/. "Ch" is usually /tʃ/ (Hard-CH) and can be /ʃ/ (Soft-CH) as in "machine," but not /s/. The letter "s" between vowels can be /z/ as in "music," but not /dʒ/.
- According to Experiment A, all of the 105 Japanese participants in the age range between 17 and 89 distinguished between [si] and [ʃi] in the Japanese contexts. The belief that Japanese speakers cannot distinguish them is not supported by this study.
- In Japanese, [e], [e:], and [ei] are all distinctive as in [e] 絵 "drawing," [e:] ええ "yes" and [ei] エイ "ray (flat fish)" while in English, there are only [ɛ] and [ei] in the mid front region; therefore, when speaking English, remember [e:] should not be used. Likewise, in Japanese, [o], [o:], and [oɯ] are all distinctive as in [ko] 子 "child," [ko:] 甲 "the back of the hand" and [koɯ][39] 請う "request" while in English, there is only

---

[38] 50 音表 "50-Sound-Chart" displays the Japanese characters, *Hiragana* or *Katakana*, in the 5×10 grid.

[39] The pair of [ko:] (the back of the hand) and [koɯ] (request) involves syllable and morpheme structures. However, both words consist of one foot (Akiyama, 2009).

[ou] in the mid back region.

- Linguistic sounds are discrete, whereas non-linguistic sounds are continuum (Hocket, 1960). Therefore, each language has its own phonological inventory, and all the words in a particular language are made of only the phonemes listed in its inventory. There are no in-between sounds in the sense of phonemes. When speaking English, always follow the English phonological system and never refer to *Katakana*-English pronunciation which is based on the Japanese phonology. English has 13 vowels and 23 consonants[40]:

| | | | | | |
|---|---|---|---|---|---|
| Long V | ā /ei/ | ē /i/ | ī /ai/ | ō /ou/ | ū /ju/ /u/ |
| Short V | ă /æ/ | ĕ /ɛ/ | ĭ /ɪ/ | ŏ /ɑ/ | ŭ /ʌ/ |
| Digraph | oi /oi/ | ou /au/ | ŏŏ /ʊ/ | | |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Voiced | b | d | v | g | | j | l | m | n | r | z | w | y | zh /ʒ/ | *th* /ð/ |
| Voiceless | p | t | f | k | h | ch | | | | | s | | | sh | th /θ/ |

***Figure 4.3.*** **The organization of English phonological inventory designed for Japanese ESL learners**

The organization in Figure 4.3 was based on Wang (1991) and School Specialty Publication (2007) with minor adjustment to a more approachable way for Japanese speakers; for example, the grouping of the English consonants based on voicing would be associated with the Japanese *Sêon* (voiceless)-*Dakuon* (voiced) distinction by Japanese speakers, as in か゛きくけこ [kakikɯkeko] vs. がぎぐげご [gagigɯgego] where the two small dots added at the top right make voiceless consonants voiced.

---

[40] Since the participants' target dialect, Canadian English, has merged /ɔ/ and /ɑ/, I did not include /ɔ/. I treated both [ə] and [ʌ] as "Short U/ʌ/" (R. Wardhaugh, personal communication, 2009). I also did not distinguish between "wh" and "w" in a sense of sound to make the inventory simple. I also did not count [ŋ] as a phoneme to make the inventory simple because whether [ŋ] is a phoneme or it is underlyingly /ng/ is debatable (R. Wardhaugh, in personal communication, 2009).

As well, the 5-vowel-letter system in English would also be associated with the Japanese 5-vowel-letter system あいうえお [aɯeo] by Japanese speakers, and the system of each of the five vowel letters having two sounds in English would be associated with the system that each Japanese *Kanji* logography has two ways to pronounce, Chinese and Japanese originated sounds.

## 4.7 Interview

At the very beginning of the lesson, I orally asked several questions aimed at gathering information about the language and linguistics background of the participants. I also gave a few quizzes to determine their knowledge of the basic English phonology and orthography. By interviewing the participants at the beginning of the lesson, I wanted to know whether the lesson would provide new, beneficial information to them. It is because the more information instructors have, the more easily instructors could assist and motivate learners (Huang, 2003). See Appendix C: Interview.

## 4.8 Results of /s/ and /ʃ/ before High Front Vowels

Table 4.2 to 4.5 show the results of the participants' production of the words that have either /s/ or /ʃ/. When the participants' production was within the acceptable range of the target phoneme, the cell in the charts is kept blank. Otherwise, the incorrectly pronounced sound is transcribed.

Table 4.2
*Pre-test: The number of errors and incorrectly pronounced sounds for words with /s/*

| | Target: /s/ 7 occasions | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | see | sea | sits | seat | CD | sick×2 | | Errors |
| P1 | | | | | | | | 0 |
| P2 | | | | | | | | 0 |
| P3 | | | ʃ | | ʃ | ʃ | ʃ | 4 |
| P4 | | s-θ | ʃ | | | | | 2 |
| P5 | | | ʃ | | | | | 1 |
| P6 | | | | | | | | 0 |
| P7 | | | s-ʃ | ʃ | ʃ | | | 3 |
| P8 | | | ʃ | | | | | 1 |
| P9 | | | | | | | | 0 |
| P10 | | | ʃ | ʃ | | | | 2 |
| P11 | | | | | | | | 0 |
| P12 | | | | | ʃ | | | 1 |
| P13 | | | | | | | | 0 |
| P14 | | | | | | | | 0 |
| P15 | | | | ʃ | | | | 1 |
| P16 | | | | | | ʃ | ʃ | 2 |
| P17 | | | ʃ | ʃ | ʃ | ʃ | ʃ | 5 |
| P18 | | s-ʃ | | | ʃ | | | 2 |
| P19 | | | ʃ | ʃ | ʃ | ʃ | ʃ | 5 |
| P20 | | | | ʃ | | | | 1 |
| P21 | | | | | ʃ | | ʃ | 2 |
| P22 | | | ʃ | ʃ | ʃ | ʃ | | 4 |
| P23 | | | | | | | | 0 |
| P24 | | | | | | | | 0 |
| P25 | | | N/A | ʃ | ʃ | | | 2 |
| P26 | | | | | | | | 0 |
| P27 | | | | | | | ʃ | 1 |
| Sum | 0 | 2 | 9 | 8 | 9 | 5 | 6 | 39 |

*Note*. s-ʃ = a sound somewhere between [s] and [ʃ]; s-θ = a sound somewhere between [s] and [θ]; N/A = not applicable because the participant pronounced a wrong word. P25 misread "sits" as "sites." It is not counted as an error because it is not a pronunciation error.

**Table 4.3**
*Pre-test: The number of errors and incorrectly pronounced sounds for words with /ʃ/*

| | **Target: /ʃ/** 11 occasions | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | washing | machines×2 | | relationship | sheets | she×6 | | | | | | Errors |
| P1 | | s | | s | s | | | | | | s | 4 |
| P2 | | s | | | s | | | | | | | 2 |
| P3 | | | | | s | | | | | | | 1 |
| P4 | | | | | | | | | | | | 0 |
| P5 | | | | | | | | | | | | 0 |
| P6 | | | | | | | | | | | | 0 |
| P7 | | s | | s | s | s | | | | | | 4 |
| P8 | | | | | | | | | | s | | 1 |
| P9 | | | | | | | | | | | | 0 |
| P10 | | | s | | | | | | | | | 1 |
| P11 | s-θ | s | s-θ | s-θ | s | s-θ | s | s | s | s | s | 11 |
| P12 | | | s | | s | | | | | | | 2 |
| P13 | | s | s-ʃ | | | | | | | | | 2 |
| P14 | s | s | s | s | s | | | | s | | | 6 |
| P15 | | | | | | | | | | | | 0 |
| P16 | s | s | s | s | s | | | s-ʃ | | | | 6 |
| P17 | | s | s | | | | | | | | | 2 |
| P18 | | | | | | | | | | | | 0 |
| P19 | | s | | | | | | | | | | 1 |
| P20 | | | | | | | | | | | | 0 |
| P21 | | s | s | s | s | s | s | s | s | s | s | 10 |
| P22 | | | | | | | | | | | s | 1 |
| P23 | | | | | | | | | | | | 0 |
| P24 | | s | s | s | s | s | s | s | s | s | s | 10 |
| P25 | | | s | | | | | | | | | 1 |
| P26 | | | | | | | | | | | | 0 |
| P27 | | s | | s | | | | | | | | 2 |
| Sum | 3 | 12 | 10 | 8 | 10 | 4 | 3 | 4 | 4 | 4 | 5 | 67 |

Table 4.4
*Post-test: The number of errors and incorrectly pronounced sounds for words with /s/*

| | Target: /s/ 7 occasions | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | see | sea | sits | seat | CD | sick×2 | | Errors |
| P1 | | | | | | | | 0 |
| P2 | | | | | | | | 0 |
| P3 | | | | | | | | 0 |
| P4 | | | | | | | | 0 |
| P5 | | | ʃ | | | | | 1 |
| P6 | | | | | | | | 0 |
| P7 | | | | | | | | 0 |
| P8 | | | | | | | | 0 |
| P9 | | | | | | | | 0 |
| P10 | | | | ʃ | | | | 1 |
| P11 | | | | **ʃ** | | | | **1** |
| P12 | | | | | | | | 0 |
| P13 | | | | | | | | 0 |
| P14 | | | | | | | | 0 |
| P15 | | | | | | | | 0 |
| P16 | | | | | | | ʃ | 1 |
| P17 | | | | | | | | 0 |
| P18 | | | | | | | | 0 |
| P19 | | | | | | | ʃ | 1 |
| P20 | | | | ʃ | | | | 1 |
| P21 | | | | | | | | 0 |
| P22 | | | | | ʃ | | | 1 |
| P23 | | | | | | | | 0 |
| P24 | | | | | | | | 0 |
| P25 | | | | | | | | 0 |
| P26 | | | | | | | | 0 |
| P27 | | | | | | | | 0 |
| Sum | 0 | 0 | 1 | 3 | 1 | 0 | 2 | 7 |

*Note*. The bold symbols indicate errors that were not committed in the pre-test. The bold numbers of error are larger than that in the pre-test.

**Table 4.5**
***Post-test: The number of errors and incorrectly pronounced sounds for words with /ʃ/***

| | **Target: /ʃ/** 11 occasions | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | washing | machines×2 | | relationship | sheets | | | she×6 | | | | Errors |
| P1 | **s-ʃ** | s-ʃ | | | | | | | | | s-ʃ | 3 |
| P2 | | | | | | | | | | | | 0 |
| P3 | | | | | | | | | | | | 0 |
| P4 | | | | | | | | | | | | 0 |
| P5 | | | | | | | | | | | | 0 |
| P6 | | | | | | | | | | | | 0 |
| P7 | | s-ʃ | | | | | | | | | | 1 |
| P8 | | | | | | | | | | | | 0 |
| P9 | | | | | | | | | | | | 0 |
| P10 | | | s | | | | | | | | | 1 |
| P11 | | s-ʃ | s | s-ʃ | s | s | s | | s | s | s | 9 |
| P12 | | | | | | | | | | | | 0 |
| P13 | | | | **s** | | | | **s** | | | | 2 |
| P14 | | | s | | | | | | | | | 1 |
| P15 | | | | | | | | | | | | 0 |
| P16 | | | | | | | | | | | | 0 |
| P17 | | | | | | | | | | | | 0 |
| P18 | | | | | | | | | | | | 0 |
| P19 | | s | | | | | | | | | | 1 |
| P20 | | | | | | | | | | | | 0 |
| P21 | | | | | | | | | | | | 0 |
| P22 | | | | | | | | | | | | 0 |
| P23 | | | | | | | | | | | | 0 |
| P24 | | | | | | | | | | | | 0 |
| P25 | | | | | | | | | | | | 0 |
| P26 | | | | | | | | | | | | 0 |
| P27 | | | | s | | | | | | | | 1 |
| Sum | 1 | 4 | 3 | 3 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 19 |

Figure 4.4 and 4.5 are the comparison between pre-test and post-test in terms of the number of errors.

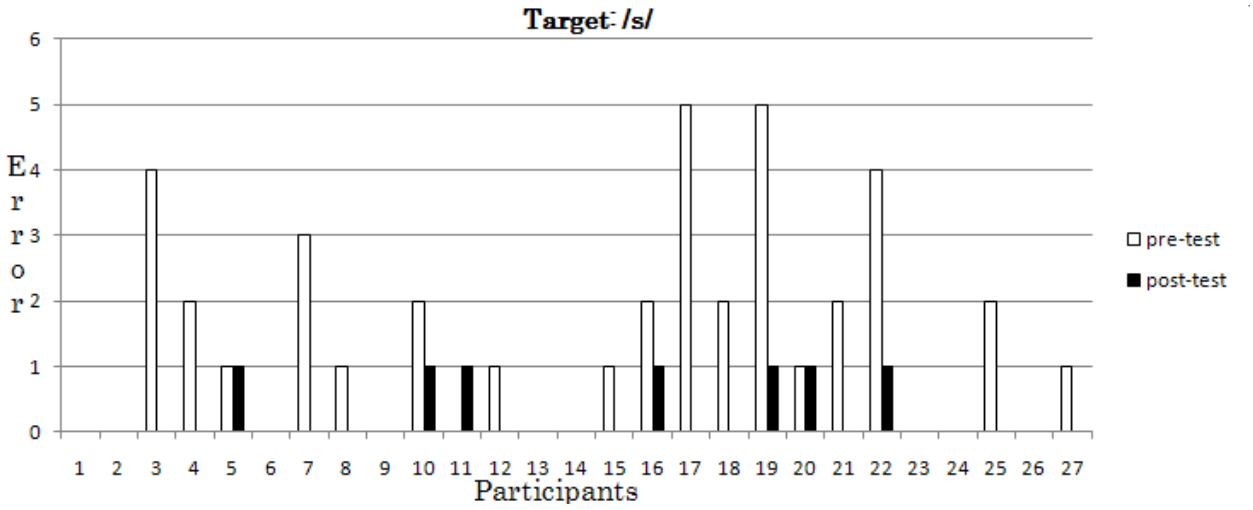


*Figure 4.4.* **The number of errors for the /s/-stimuli in pre-test and post-test**

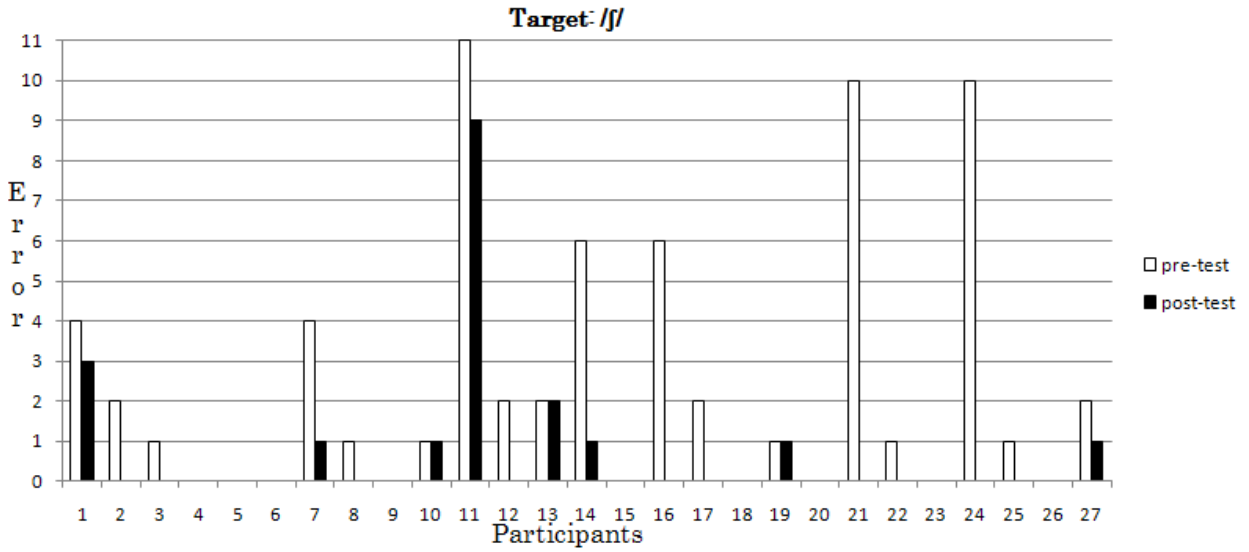


*Figure 4.5.* **The number of errors for the /ʃ/-stimuli in pre-test and post-test**

The results show that the difference between pre-test and post-test analyzed by two-tailed paired t-test is significant ($p < 0.0005$ in the [s]-stimuli and $p < 0.005$ in the [ʃ]-stimuli). Although they did not receive any advice of articulation, but took mere 40-minute

advice about the concept of phoneme and the spelling rules, the participants corrected their pronunciation immediately. This indicates that their mispronunciation regarding [s] and [ʃ] is not derived from an inability to articulate, but from a lack of the knowledge needed to produce them correctly.

In the pre-test, nine participants pronounced all the occurrences of /s/ correctly, and also nine participants pronounced all the occurrences of /ʃ/ correctly. Only P6, P9, P23, and P26 correctly pronounced both /s/ and /ʃ/ on all the occasions. As I will mention in 4.10, P6 and P23 had an experience of being taught English by native English-speaking teachers in Japan and P9 had been partially taught IPA in Japan. P26 was not taught English by native English-speaking teachers in Japan, but he enthusiastically took many pronunciation classes at the ESL school and asked his Canadian friends to correct his pronunciation during the eight months of residence in Canada. The other participants made at least one error. Therefore, 23 out of the 27 participants (85%) pronounced /s/ and/or /ʃ/ incorrectly, which matches Fukushima et al.'s (2008) study that showed over 80% of Japanese learners of English cannot correctly distinguish between /s/ and /ʃ/ before high front vowels. Moreover, when I started to talk about /s/ and /ʃ/ in the lessen, some participants, such as P11, P25, and P27, reported that they were not comfortable with /s/ and /ʃ/ and they wanted some tips.

According to Table 4.2 and 4.3, some words are more likely to be mispronounced than others. Specifically, "see" and "sea" were unexpectedly pronounced well. For some reason, "s" in "sits" was one of the [s]'s that were most often mispronounced as [ʃ], tying "CD." "Seat" was also mispronounced frequently. Since "CD" and "seat" were imported to the Japanese vocabulary as [ʃi:di:] シーディー and [ʃi:to] シート respectively, the errors in

these words can be considered as loanword interference. As for /ʃ/, "sh" in "washing" was most often correctly pronounced. Since this "sh" is before an inflectional morpheme "-ing," the [ʃ] + [ɪ] sequence in "washing" is not as strongly connected as that in the other words because of the presence of the morpheme boundary. In fact, Eckman, Elreyes, and Iverson (2003) predicted the stages of development for Korean ESL learners whose L1 has [s] and [ʃ] as allophonic variations but not as two different phonemes; Korean learners acquire this contrast in derived words, such as "messing" and "meshing," earlier than nonderived words, such as "sea" and "she" (as cited in Gass & Selinker, 2009). On the other hand, "ch" in "machine" and "sh" in "sheets" were often mispronounced as [s]. These words were also imported to the Japanese vocabulary as [maʃi:N] マシーン and [ʃi:to] シート. These errors seemed to have come from over-generalization that [ʃi] is more Japanese-like and [si] is more exotic. This is considered to be a type of "avoidance" of L2 acquisition. Gass and Selinker (2009) stated that when learners find great similarities between L1 and L2, the learner may doubt that these similarities are real. As mentioned in 2.1, since in Japanese [ʃ] before high front vowels occurs in native Japanese vocabulary while it may be still safe to consider [s] before high front vowels as foreignism, some participants doubted that such an L1-like sound structure could occur in L2, and mis-generalized that [s] before high front vowels were more L2-like. Makino (2008b) called Japanese learners' depalatalization "palatalization hypercorrected" (p.27). Interestingly, all the participants who mispronounced "sheet" as [si:t] correctly pronounced "seat," except for P7, while all the participants who mispronounced "seat" as [ʃi:t] correctly pronounced "sheet," except for P7. Therefore, these participants, except for P7, may not have realized "seat" and "sheet" as different entities as both "seat" and "sheet" are adapted to [ʃi:to] in the Japanese

vocabulary. More intriguingly, the words "sick," "machine," and "she," which occur more than once in the passage, were not consistently pronounced in the same way by the same speaker. For example, by seven participants, "ch" in "machine" was pronounced as [ʃ] at one occasion and as [s] at the other. This inconsistency implies that the participants were not sure of the proper target sound.

Another finding is that some participants made errors only in either [ʃ] or [s].

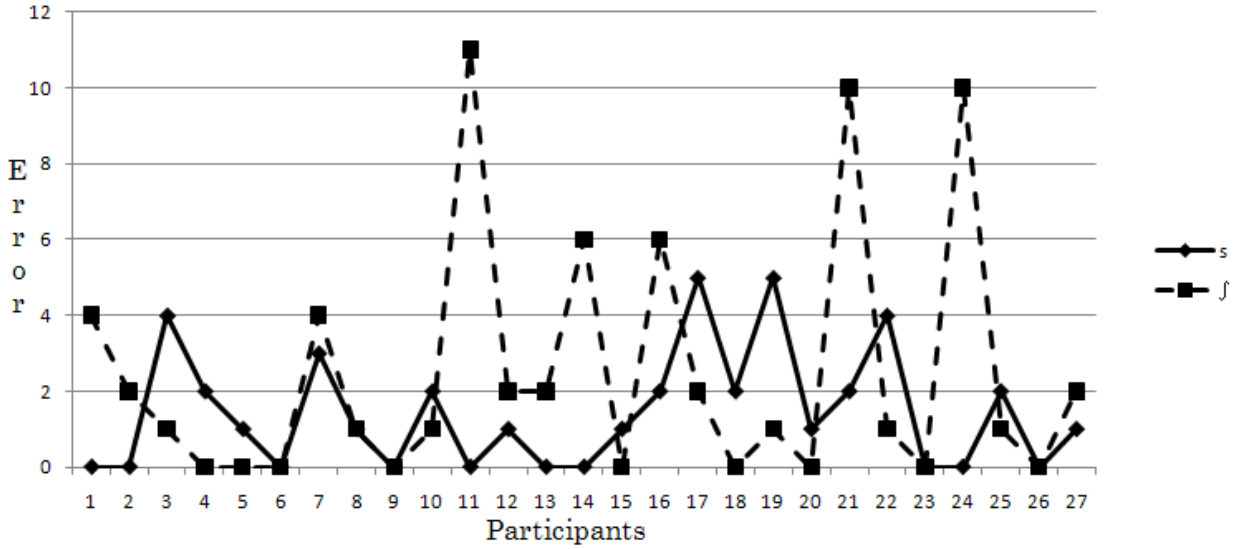


***Figure 4.6.*** **The number of error by each participant for the [s]-stimuli and the [ʃ]-stimuli in pre-test**

According to Figure 4.6, for example, P1, P2, P11, P13, P14, and P24 made quite a few errors with [ʃ] but made no errors with [s]. In contrast, P3, P4, P18, P19, and P22 notably performed better with [ʃ]. Those participants were biased toward either [s] or [ʃ]. The extreme case is P11, who pronounced all the [s]'s before high front vowels correctly while all the [ʃ]'s before high front vowels incorrectly. P24 showed almost the same characteristic. The opposite extreme was not found. I consider that the bias toward [s] to be avoidance of a similarity to L1, while the bias toward [ʃ] is considered to be loanword interference. Some participants, such as P7, P10, P12, P25, and P27 did not show a particular bias toward one

of the two, but randomly made errors. It is not clear whether there is a relationship among variables.

In the post-test, 15 out of the 27 participants did not make even a single mistake, and 22 participants made at most only one error in total among both [s]-stimuli and [ʃ]-stimuli. The number of the total errors regarding [s] and [ʃ] by all the participants dropped from 106 to 26. The average number of errors by each participant dropped from 3.93 to 0.96. This improvement is quite remarkable.

P5, P13, and P20 did not show improvement regarding /s/ and /ʃ/. However, firstly, these three performed fairly well even in the pre-test, and it is possible to make some nervous mistakes in an experimental environment. Secondly, the /ei/-stimuli, the /ou/-stimuli, and the interview show their improvement (See 4.10 and 4.12). As for P11, she made no errors regarding [s] in the pre-test but made one error in the post-test; namely, she pronounced "seat" correctly in the pre-test but pronounced it as [ʃit] in the post-test. However, this regression can actually be considered as progress in the process of "U-shaped learning"; at this stage, learners discern some rules in TL (target language) but cannot utilize the rules correctly in the TL standards (Brown, 2007). In fact, in the pre-test, she replaced all the occasions of [ʃ] before high front vowels with [s] with no exceptions. This is aforementioned "avoidance"; L2 learners may doubt that similarities between L1 and L2 are real. Actually, when I explained that "machine" was not [məsin], but [məʃin], some of the participants, including P11, found it difficult to accept it and kept pronouncing [məsin] until I said that "chi" in "machine" was supposed to be like [ʃiː] in the Japanese word [ʃiːtake] 椎茸 "shitake mushroom." However, in the post-test, P11 finally became determined to use [ʃ] before high front vowels even though she did not necessarily use it at

a proper place. That can be considered as a progress in the psychological aspect.

### 4.9 Phonetic Aspects of /si/ and /ʃi/

Firstly, the most interesting finding is that none of the participants in this experiment pronounced /si/ as [swi]. In other words, every participant pronounced pure [si] without on-glide in the English contexts. Recall that in Experiment A in Chapter Three, 46 out of 93 in JG1 and six out of 12 in JG2 pronounced [swi] in the Japanese contexts. Since the participants in Experiment B are only in their 10's, 20's, and 30's, we should turn our attention to the participants in the same generations in Experiment A. Still, 10 out of 35 in JG1 and 5 out of 10 in JG2 pronounced [swi], which means that one third of the younger generations tend to add on-glide in the Japanese contexts. Moreover, in another study (Nogita, 2010), four female Japanese ESL learners in the age range 19 and 33 pronounced both Japanese and English stimuli including the [si] syllables. The results showed that two of the four participants, the 19-year-old and the 33-year-old, pronounced [swi] in the Japanese contexts while all of them pronounced [si] without on-glide in the English contexts. Considering this asymmetry, the contrast between the English /s/ and /sw/, as in "seat" and "sweet," does not seem problematic for Japanese ESL learners, contrary to my concern. Probably, the acquisition of [si] without on-glide is easy, or some Japanese speakers may consciously choose whether they insert on-glide depending on the contexts.

Secondly, the quality of [ʃ] was interesting as well. In the pre-test, P1, P7, P11, P13, P15, P16, P18, and P25 occasionally pronounced something between [s] and [ʃ], which sometimes made it difficult for the judges to transcribe their production. As I explained in 3.3.8, some Japanese speakers, especially young women, sometimes pronounce [s]-like [ʃ] even in the English contexts. These participants may occasionally pronounce [s]-like [ʃ] in

Japanese, and they transferred this habit to English. However, in the post-test, those participants, except P1 and P11, pronounced clear [ʃ]. P7 pronounced [s]-like [ʃ] once in the post-test, but in the other occasions, she corrected herself when she happened to pronounce [s]-like [ʃ]. As for P15, he seemed to emphasize the difference between [s] and [ʃ], and his [ʃ] sounded somewhat closer to the retroflex [ʂ]. These participants seem to be sensitive to phonetic differences. In contrast, P1 and P11 still pronounced [s]-like [ʃ] in the post-test. P1 pronounced /ʃ/ as [sʃ] three times in the post-test, but the [ʃ] part was too short to be perceived as Eng/ʃ/. Therefore, those three productions were counted as errors. Recall that some Japanese speakers tend to pronounce [sʃ] for pre-pausal /ʃ/ as shown in Figure 3.4. Whether P1 intended to pronounce /ʃ/ correctly or she still misconceived /ʃ/ as /s/ was not clear. To make matters more complicated, P1 did not feel well during the lesson, so it is also possible that she could not concentrate on clear pronunciation. As for P11, she showed most interest in this experiment and asked many questions after the experiment. Although she performed worst in the post-test in terms of score, she tried the hardest. She may have mulled over the symbol-sound correspondence rules so much that she could not afford to think of phonetic aspects and kept pronouncing [s]-like [ʃ].

P10 and P24 showed their sensitivity to the phonetic aspects as well. In the pre-test, P10's /ʃ/ showed some phonetic variations. She pronounced English-like [ʃ] a couple of times, but she mostly pronounced either the prototypical Japanese [ɕ] or more retroflex [ʂ] which appears neither in L1 nor L2. On the other hand, in the post-test, most of her productions of /ʃ/ before high front vowels were retroflex [ʂ]. Since [ɕ] is produced in a more back region than [ʃ] which is produced in a more back region than [ʂ], her productions can be analyzed as over-correction. As mentioned above, P15 showed a similar characteristic.

As for P24, she quite consistently pronounced the alveolo-palatal [ɕ], which is more L1-like, for /ʃ/ in both pre-test and post-test, but her productions were categorized as Eng/ʃ/ by the judges. Interestingly, during the lesson, she, herself commented that her Jap/ʃ/ sometimes sounded like [s]. This means that she perhaps had the habit of pronouncing [s]-like [ʃ] and she was aware of that. In fact, just like P7 did, P24 also corrected herself when she happened to pronounce [s]-like [ʃ] during recording in the post-test. Except for the particular cases of P1 and P11, Japanese ESL learners seem quite sensitive to the phonetic differences between L1 and L2 in terms of these coronal fricatives.

Thirdly, [θ]-like [s] was occasionally produced by P4 and P11. In the pre-test, P4 pronounced "sea" as [θi:], but she seemed to intend to pronounce [si:] and the [s] happened to be more dental, just as in Experiment A (see Table 3.3), a few participants occasionally pronounced more dental [s] than the prototype [s]. As I mentioned in 2.6, Jap/s/ is less sibilant and possibly more dental than Eng/s/ (Beckman et al., 2009) and probably Jap/s/ appears as [θ] as a weakened form (Inozuka et al., 2009). Therefore, I interpret that P4 happened to pronounce the weakened form of Jap/s/ in the pre-test. In fact, in the post-test, she always pronounced clearer [s] since she paid more attention to /s/ and /ʃ/. Moreover, in terms of consonant qualities, P4 performed the best among all the participants; she pronounced all the three occasions of /θ/ in "think," "thing," and "healthy" correctly and were also good with [l], [ɹ], [ð], and [v], which are said to be problematic consonants for Japanese ESL/EFL learners, compared with any of the other participants. P4 seemed fairly flexible about consonant qualities. As for P11, she also pronounced something between [s] and [θ] a few times when the target was either /s/ or /ʃ/ in the pre-test. P11's case is also considered as the weakened form of Jap/s/; she did not seem to attempt the wrong target /θ/

for /s/ or /ʃ/ because she almost always substituted /θ/ with a different consonant, which means that P11 did not own the phoneme /θ/ in the inventory of her inter-language (IL), unlike P4. In the post-test, she still pronounced somewhat [θ]-like [s] but overall her [s] became clearer than in the pre-test. Overall, the participants quite flexibly tried to adjust their consonant quality in their ILs in the sounds such as [sw], [s], and [ʃ].

## 4.10 Results of Interview

The oral interview was designed to examine the participants' awareness about the English phonological and orthographic systems and their correspondence.

### *4.10.1 Questions Regarding the Participants' Background*

**a.** "Have you ever been taught IPA in Japan? Can you actually utilize it?"

Only P9 knew IPA to some extent. The others answered "no." P9 was taught it in high school for the university entrance exams, not for perception and production. However, she was one of the four participants who did not make any errors regarding /s/ and /ʃ/ in both pre-test and post-test. A few other participants were partially taught IPA in high school, but they forgot it.

**b.** "Have you ever been taught any of "phonics," "42-sounds[41]," "learn to read," or an equivalent?"

Only P23 had been partially taught phonics, but not quite systematically. The others answered "no." In Japan, P23 was working for a kindergarten where half of the students were American. Therefore, P23 was taught phonics for her job, but was not taught in English classes in high school. She was one of the four participants who did not make

---

[41] There seem to be some versions of "42-sounds." According to Jolly Phonics, the 42 sounds are s, a, t, i, p, n, c k, e, h, r, m, d, g, o, u, l, f, b, ai, j, oa, ie, ee, or, z, w, ng, v, *oo* [ʊ], oo [u], y, x, ch, sh, *th* [ð], th [θ], qu, ou, oi, ue, er, and ar.

any errors regarding /s/ and /ʃ/ in both pre-test and post-test. P23 was one of the only two participants who distinguished [l] and [ɹ], but she had not yet acquired the voiced fricatives [v], [ʒ], or [z] and her vowels sound mostly based on *Katakana*-English. A few other participants had heard of phonics, but they did not know what it was at all.

**c.** "Have you ever been taught proper English pronunciation in Japan?"

P4, P6, P11, and P23 had. P4's and P6's majors in their university in Japan were related to international communication. P11 was taking English lessons other than school classes in Japan. P23 was a teacher in a kindergarten where half of the students were American. Note that P6 and P23 were two of the four participants who did not make any errors regarding /s/ and /ʃ/ in both pre-test and post-test. Moreover, P4 and P23 were the only participants who distinguished between [l] and [ɹ]. P11 did not perform well in terms of the tasks of this experiment, but she was most interested in pronunciation. What is more important is that those four participants commented that they were taught each phoneme, but were not taught the whole picture of the English phonological inventory.

*4.10.2 Quizzes*

**a.** "Can you identify whether the following words in each pair are the same or different in pronunciation? If different, can you identify the difference?"

"meat" versus "meet," "ear" versus "year," "bone" versus "born," "cars" versus "cards," "branch" versus "brunch"

None of the participants, except for P26, knew the answers, not even P9, who had been partially taught IPA, P23, who had been partially taught phonics, P6 and P11, who had been taught pronunciation by native English-speaking teachers, and P13, who was a regular university student in Canada. P17 happened to know that "meat" and "meet" were

homophones, but she knew only this particular pair and was not aware of the phonological structure of [mit]. She did not know that "meat" and "meet" rhyme with words like "eat" and "Pete," either. As for P26, he said that he could not pronounce these words properly but he knew that "meat" and "meet" were homophones and the other pairs were not homophones. He also identified the differences between "ear" and "year," "bone" and "born," "cars" and "cards," and "branch" and "brunch." P26 was the only one who answered this question correctly. Reportedly, P26 was taught these words in pronunciation classes in his ESL school in Canada, not in Japan.

**b.** "Do you know that in English, unstressed vowels are reduced to Short U [ə]? E.g. "J<u>a</u>pan," "comm<u>e</u>nt," "terr<u>i</u>ble," "comm<u>o</u>n" and so on."

None of the participants knew this rule. They seemed to find it difficult to accept this rule. Some of the participants recalled that they had been taught it in an ESL school, but they did not quite understand the concept. Since the English schwa is adapted to various Japanese vowel categories depending on the word in the loanword adaptation process (Suzuki, 2008; Horiguchi, 1997), Japanese learners of English have no way to acquire the schwa rule from *Katakana*-English.

**c.** "For the word "idea," which do you think sounds more proper between [aɪdiə] and [aɪdiɹ]?"

Interestingly, all the participants answered that [aɪdiɹ] was proper and [aɪdiə] was wrong. They all had a stereotype that the more rhotacized, the more English-like. In other words, [ɹ] may not be a linguistic consonant phoneme, but just "spice" for an "English-like" flavour. This misconception must be one of the complexities of the well-cited [l]-and-[ɹ] issue of Japanese ESL learners. I can see a lack of their phonological and orthographic

awareness here again. Note that Japanese ESL learners' [ɹ] insertion is derived from a completely different motivation from that in British English, which is for avoidance of a vowel sequence.

*4.10.3 Question for the Participants' Opinion*

> "Do you agree with the process, *Katakana*-English-based underlying representation → Stereotypical L2 flavours (e.g. "The more tongue rolling[42] (rhotacized), the more English-like.") → Surface representation in IL?"

All the participants strongly agreed. Only P9, who had been partially taught IPA, commented that she partially changed the *Katakana*-English-based underlying representation (UR). For example, when the *Katakana*-English-based UR is /ɾi:do/ "read," P9 changed it to /ɹi:do/ by using her IL phoneme /ɹ/. Otherwise, she used the same process. P13, a regular university student, commented that her UR was not only based on *Katakana*-English, but also her UR can be the English words read in the *Rômaji* (Japanese Romanization) ways. After all, none of the participants had proper knowledge of the target sounds in English.

*4.10.4 Discussion of the Interview*

This interview yielded the following findings: in English, the participants were conscious that they lacked the knowledge of the connections between orthography and the sounds, and they may have had partial knowledge of the English sound system but lacked the whole picture.

Firstly, when I explained that every word consists of phonemes in the phoneme inventory in the particular language, the participants all agreed that the English words

---

[42] Recall that the lesson was done in Japanese. "舌を巻く" or "巻き舌" is the expression I actually used in the lesson for "tongue rolling."

produced by them did not necessarily consist of phonemes, because they had not been formally taught the whole picture of the English phoneme inventory. In other words, they could not phonemically segment English sounds. Some of the participants, such as P4, P9, P23, and P26 seemed to have the concept of phoneme for some particular consonants, while none of the participants seemed to completely consider the vowels as discrete units. Interestingly, although all the participants' vowels sounded mostly based on *Katakana*-English, I cannot say that they completely transferred L1 "phonological categories" to L2. For example, no one was sure that "meet" and "meat" were homophones and many of the participants suspected that these two words could be slightly different, except P17 and P26 happened to know this particular pair. "Slightly different" implies that they cannot divide "meet" and "meat" into phonemes or any phonological units because categorization of phonological units always has clear-cut boundaries and does not have the concept of slight differences. Importantly, all the three English phonemes in /mit/ are expected to be categorized as Jap/m/, Jap/iː/, and Jap/t/ respectively, except phonotactically /t/ does not appear in this position in standard Japanese. Moreover, both "meat" and "meet" were imported to the Japanese vocabulary as homophones including the accent pattern, [miↄ↓ːto] (the arrow shows the accent pattern). Strangely, even when L1 has similar counterpart phonemes, the participants still may have perceived [mit] as non-linguistic continuum sounds. Therefore, Japanese ESL learners may often recognize English sounds as non-linguistic continuum sounds while referring to *Katakana*-English pronunciation.

Secondly, the participants all insisted that they made no connections at all between the orthography and the sounds in English, even P4, P6, and P11, who had been taught English by native speaking teachers in Japan, P9, who had been partially taught IPA, and P23, who had been partially taught phonics. In fact, the word pairs, "ear" versus "year,"

"bone" versus "born," "cars" versus "cards," and "branch" versus "brunch," obviously show the phonemic differences on the spellings, but the participants, except for P26, had no idea of the sounds. This indicates a lack of their awareness of orthographical information. Since orthographic knowledge precedes phoneme awareness at the conscious level (Carroll, 2004), it is understandable that the participants lacked phoneme awareness.

Thirdly, the participants all agreed that they did not have the whole picture of the English phonological inventory. For example, P9 had been taught that "bat" and "but" were /bæt/ and /bʌt/ with IPA, so she knew this particular minimal pair. However, she could not at all apply this knowledge to the same type of minimal pairs, such as "branch" versus "brunch," or "yam" versus "yum." As well, she was not aware that "a" in "bat" was phonologically the same as "a" in "apple" and "u" in "but" was the same as "u" in "under." P23's knowledge of phonics showed the same tendency. P26 also showed the same tendency. The other participants also claimed that they were taught the pronunciation of some words in ESL classes, but they could not apply the knowledge to other words. This can be understood as that since they lacked the idea of discreteness of phoneme in English, they were not able to figure out the English phonological grammars. Makino (2009a) suggested teaching all the English phonemes to show the whole picture of the inventory to Japanese EFL learners at the early stage. Therefore, the charts like Figure 4.3 looked helpful for the participants because the charts enabled the participants to understand the general picture of the English phonological inventory and the idea of discreteness of phoneme in English.

In the pre-test, the participants seemed to lack the phonological awareness and orthographical knowledge in English, and they were conscious of what was missing in their IL. At this rate, it is possible that many Japanese ESL learners have no clear-cut phoneme boundary between [s] and [ʃ] in the English contexts, although in the Japanese contexts,

they do have a clear-cut boundary between [s] and [ʃ]. At the same time, there is no wonder those who lack the concept of phonological categorizations are flexible about phonetic differences.

### 4.11 Results of /z/

In this paper, I did not focus on the voiced counterpart of Eng/s/, namely Eng/z/, because the case is more complex than /s/. In most dialects in modern Japanese, the pure fricative [z] is not a phoneme, but is considered as a weakened form of the voiced affricate /dz/ (Grenon, 2005). However, I still observed two occasions of [z] in the stimulus, namely "music" and "easily." The [z] in these words is inter-vocalic and in an unaccented syllable, so it is expected that Japanese speakers fricate them as they often fricate intervocalic Jap/dz/ in an unaccented syllable in Japanese. As well, [z] in "music" is before a high front vowel and I expected that some of the participants would pronounce it as [ʒ] or [dʒ] being influenced by the loanword /mjɯːdʒikkɯ/ ミュージック "music." Just like [dz] and [z], in most dialects in modern Japanese, [ʒ] and [dʒ] are also not phonemic, but are allophonic variations. The fricative [z] in "easily" is not before a high front vowel, but I expected that some of the participants might have been influenced by one of the possible *Katakana*-English ways, [iː(d)ʒiɾiː], and would pronounce this /z/ as [ʒ] or [dʒ]. In this experiment, I did not count affricated /z/ as an error because to produce a pure voiced fricative may require some articulatory training and the purpose of this experiment was to find out whether orthographic and phonological knowledge improves Japanese ESL learners' pronunciation. Therefore, only when /z/ was palatalized to [ʒ] or [dʒ], it was counted as an error. When the participants pronounced [dz], I parenthesized the symbols in Table 4.6 and 4.7.

**Table 4.6**
*Pre-test for /z/*

| Target:/z/ 2occasions | | | |
|---|---|---|---|
| | mu<u>s</u>ic | ea<u>s</u>ily | Error |
| P1 | | | 0 |
| P2 | | | 0 |
| P3 | ʒ | | 1 |
| P4 | | (dz) | 0 |
| P5 | | | 0 |
| P6 | | | 0 |
| P7 | ʒ | (dz) | 1 |
| P8 | | | 0 |
| P9 | | | 0 |
| P10 | ʒ | dʒ | 2 |
| P11 | dʒ | | 1 |
| P12 | | | 0 |
| P13 | | | 0 |
| P14 | | | 0 |
| P15 | ʒ | | 1 |
| P16 | | | 0 |
| P17 | | | 0 |
| P18 | dʒ | (dz) | 1 |
| P19 | | (dz) | 0 |
| P20 | | | 0 |
| P21 | ʒ | | 1 |
| P22 | ʒ | | 1 |
| P23 | | (dz) | 0 |
| P24 | | (dz) | 0 |
| P25 | ʒ | | 1 |
| P26 | | | 0 |
| P27 | | | 0 |
| Sum | 9 | 1 | 10 |

**Table 4.7**
*Post-test for /z/*

| Target: /z/ 2 occasions | | | |
|---|---|---|---|
| | mu<u>s</u>ic | ea<u>s</u>ily | Error |
| P1 | | | 0 |
| P2 | | | 0 |
| P3 | | | 0 |
| P4 | (dz) | (dz) | 0 |
| P5 | | | 0 |
| P6 | | | 0 |
| P7 | (dz) | (dz) | 0 |
| P8 | | | 0 |
| P9 | | | 0 |
| P10 | | | 0 |
| P11 | ʒ | | 1 |
| P12 | | | 0 |
| P13 | | | 0 |
| P14 | | | 0 |
| P15 | | | 0 |
| P16 | | | 0 |
| P17 | | | 0 |
| P18 | (dz) | | 0 |
| P19 | | (dz) | 0 |
| P20 | | | 0 |
| P21 | | | 0 |
| P22 | | | 0 |
| P23 | | | 0 |
| P24 | | (dz) | 0 |
| P25 | (dz) | | 0 |
| P26 | | | 0 |
| P27 | | | 0 |
| Sum | 1 | 0 | 1 |

In the pre-test, nine participants pronounced [ʒ] or [dʒ] for the /z/ in "music" and one did so for the /z/ in "easily." As I expected, since "music" is a part of Japanese vocabulary as /mjɯːdʒikkɯ/, the loanword pronunciation influenced the participants' production. "Easily" is not a part of Japanese vocabulary, and the participants were not much influenced by the *Katakana*-English pronunciation. In the post-test, only one participant still palatalized the /s/ in "music," but all the others pronounced both "music" and "easily" correctly. That indicates that when the Japanese ESL/EFL learners in my study palatalize both English voiced and voiceless alveolar fricatives, it is not an issue of articulation, but an issue of a lack of knowledge.

### 4.12 Results of /ei/ and /ou/

In order to substantiate the claim that the [s] and [ʃ] confusion by Japanese ESL/EFL learners is not an issue of articulation but an issue of lack of knowledge, I also observed their /ei/ and /ou/. Japanese phonotactically allows any two vowels to be adjacent to each other, although some sequences are rare. As well, Japanese allows extra-long syllables with a complex nucleus, as in /waiN/ ワイン "wine" and /hait.ta/[43] 入った "entered" (Vance, 2008). Thus, the diphthongs Eng/ei/ and Eng/ou/ are expected to be acquired by Japanese speakers without being blocked by L1 constraints. We will now discuss the results of /ei/ first.

---

[43] Japanese has the vowel sequence $V_1V_2$ where $V_2$ is a high vowel. It is debatable whether this combination is a diphthong within the same syllable or two separate syllables. It is because this issue involves some complexities, such as accent patterns and morphological structures (Vance, 2008).

Table 4.8
*Pre-test: The number of errors and incorrectly pronounced sounds for words with /ei/*

| Target: /ei/ 12 occasions | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | stay | Canadian | able | pay | they | Canadian | name | relation | day | paint | paper | arrange | Error |
| P1 | | a | ɛ | | | a | | | | | | ɛ | 4 |
| P2 | | a-ʌ | ɛɹ | | | a | | | | | ɛ | ɛ | 5 |
| P3 | | a | ɛ | | e.i[a] | a | | | | | ɛ | ɛ | 6 |
| P4 | N/A | a | ɛ | | | a | | | | | | ɛ | 4 |
| P5 | ɛ | a | | | | a | | ɛ | | | ɛ-e | ɛ | 6 |
| P6 | | a-ɛ | ɛ | | | a-ɛ | | | | | | ɛ | 4 |
| P7 | | a-ɛ | | | | a-ɛ | | | | | | ɛ | 3 |
| P8 | ɛ | a | | ɛ | | ɑ | | ɛ | | | | ɛ | 6 |
| P9 | | a | | | | a | | | | | | ɛ | 3 |
| P10 | | a | ɛ | | | a | | | | | | ɛ | 4 |
| P11 | ɛ | ə | ɛ | ɛ | | ə | ɛ | ɛ | ɛ | ɛ | ɛ | ɛ | 11 |
| P12 | N/A | a | | ɛ | | a | | | | | | ɛ | 4 |
| P13 | | ɛ | | | | | | ə | | | | ɛ | 3 |
| P14 | | a | ɛ-e | | | a | ɛ-e | | | | | ɛ | 5 |
| P15 | | a | | | | a | | | | | | ɛ | 3 |
| P16 | | N/A | | | | a | ɛ | | | | | ɛ | 3 |
| P17 | | ʌ | | ɛ | | ʌ | ɛ | ɛ | | | ɛ | ɛ | 7 |
| P18 | | ɑ | ɛ | | | ɑ | ɛ | ɛ | ɛ | | ɛ-e | ɛ | 8 |
| P19 | | ʌ | | | | a | | | | | | ɛ | 3 |
| P20 | | a | | | | a | | | | | | ɛ | 3 |
| P21 | | a | ɛ | | | a | | | | | | ɛ | 4 |
| P22 | ɛ | a | ɛ | | N/A | a | | | | | ɛ | ɛ | 6 |
| P23 | | a | | | | ɑ | | | | | | ɛ | 3 |
| P24 | | a | ɛ | | | ɛ | ɛ | | | | | ɛ | 5 |
| P25 | | ə | ɛ | | | ə | | | | | | ɛ | 4 |
| P26 | ɛ | ɛ | ɛ | | | ɛ | ɛ | ɛ | | | | ɛ | 7 |
| P27 | ɛ | ə | ɛ | | | a | | | | | | ɛ | 5 |
| Sum | 6 | 26 | 15 | 4 | 1 | 26 | 7 | 7 | 2 | 1 | 7 | 27 | 129 |

*Note.* N/A = not applicable because the participants pronounced a wrong word. For example, he or she misread "stay" as "study." It is not counted as an error because it is not a pronunciation error.
[e.i][a] pronounced by P3 sounded like two syllables, instead of a diphthong.

Table 4.9
*Post-test: The number of errors and incorrectly pronounced sounds for words with /ei/*

| | **Target: /ei/** 12 occasions | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | stay | Cana-dian | able | pay | they | Cana-dian | name | rela-tion | day | paint | pa-per | ar-range | Er-ror |
| P1 | | | | | | | | | | | | | 0 |
| P2 | | | | | | | | | | | | | 0 |
| P3 | | | | | | | | | | | | | 0 |
| P4 | | | | | | | | | | | | | 0 |
| P5 | | | | | | | | | | | | | 0 |
| P6 | **ɛ** | | | | | | | | | | | | 1 |
| P7 | | | | | | | | | | | | | 0 |
| P8 | ɛn | | | | | | | | | | | | 1 |
| P9 | | | | | | | | | | | | | 0 |
| P10 | | | | | | | | | | | | | 0 |
| P11 | | | ɛ | | | | | ɛ | | ɛ | | ɛ | 4 |
| P12 | | | | | | | | | | | | | 0 |
| P13 | | | | | | | | | | | | | 0 |
| P14 | | | | | | | | | | | | | 0 |
| P15 | | | | | | | | | | | | | 0 |
| P16 | | | | | | | | | | | | ɛ | 1 |
| P17 | | | | | | | | | | | | | 0 |
| P18 | | | | | | | | | | | | ɛ | 1 |
| P19 | | | | | | | | | | | | | 0 |
| P20 | | | | | | | | | | | | | 0 |
| P21 | | | | | | | | | | | | ɛ | 1 |
| P22 | | | | | | | | | | | ɛ | | 1 |
| P23 | | ɛ | | | | ɛ | | | | | | | 2 |
| P24 | | ɛ | | | | ɛ | | | | | | ɛ | 3 |
| P25 | | | | | | | | | | | | | 0 |
| P26 | | | ɛ | | | | | | | | | | 1 |
| P27 | | | | | | | | | | | | | 0 |
| Sum | 2 | 2 | 2 | 0 | 0 | 2 | 0 | 1 | 0 | 1 | 1 | 5 | 16 |

*Note*. The bold symbols indicate errors that were not committed in the pre-test.

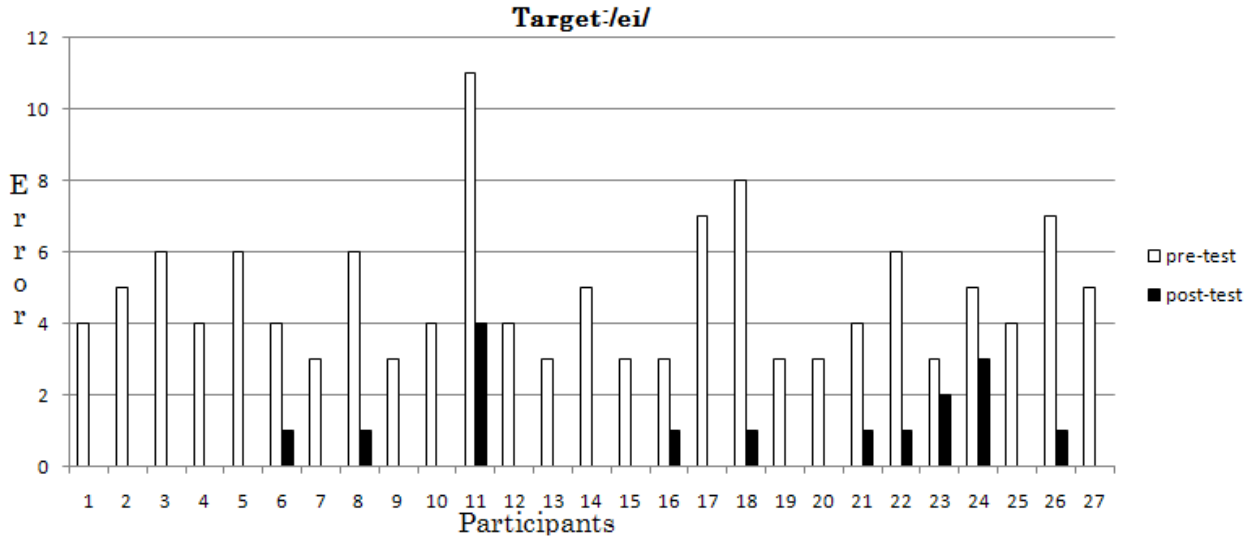


***Figure 4.7.*** **The number of errors for the /ei/-stimuli in pre-test and post-test**

Between the pre-test and the post-test, the average number of errors by each participant dropped from 4.78 to 0.59. The change from the pre-test to the post-test is notable. In the pre-test, comparing with the consonants, the qualities of the participants' vowels varied in a wider range. It is highly likely because the participants lacked the phonological awareness in English vowels even more than they did so in consonants. Moreover, there is a tendency of error, just like in the /s/ and /ʃ/-stimuli. In the /ei/-stimuli, all of the participants pronounced /ei/ in "arrange" as [ɛ] or the like, which is categorized as Eng/ɛ/, and all, except for P13 who had been in an English-speaking country for the longest time among the participants, pronounced /ei/ in "Canadian" as [a] or a similar vowel. This can be considered as loanword interference, namely /aɾeNdʒi/ アレンジ "arrange" and /kanadiaN/ カナディアン "Canadian." Note that the Japanese mid front vowel /e/[44] is not a diphthong. In the words "paper" and "able," it seems some of them tried to make a

[44] The Japanese short monophthong /e/ is between the cardinal vowels [e] and [ɛ], so it can be transcribed as [ẹ] (Inozula et al., 2009).

distinction from Eng/ɛ/ by using the Japanese long vowel /e:/. In fact, there is the English loanword /pe:pa:/ ペーパー "paper" in Japanese, in which Eng/ei/ was adapted to Jap/e:/. However, Jap/e:/ in their production sounded still closer to Eng/ɛ/. One of the possibilities is that the participants tried to change their rhythm pattern from the Japanese mora-timed rhythm to the English stress-timed rhythm consciously or unconsciously, which often made their vowel length distinction less clear than in the Japanese contexts. Since sentence level rhythm patterns are different from phonemic categorization, it is possible that those who lack the idea of discreteness of phoneme in English are sensitive to the rhythm patterns. According to Lin (2008), Mandarin speakers, whose L1 is a syllable-timed language, can master the vocalic duration patterns of English more easily than intervocalic duration patterns. Since Mandarin and Japanese show similar values of the proportion of vocalic interval within a sentence, Japanese speakers also may be able to acquire the English vocalic duration patterns without much difficulty. Interestingly, "paint" was correctly pronounced by all but P11. It can be because in the English loanword /peiNto/ ペイント "paint," Eng/ei/ is adapted to the diphthong (or the vowel sequence) Jap/ei/. This happens to be a positive L1 transfer. To explore this further, in the loanword adaptation processes, Eng/ei/ is often adapted to the long vowel Jap/e:/ as in /pe:pa:/ ペーパー "paper," and sometimes to the diphthong (or the vowel sequence) Jap/ei/ as in /peiNto/ ペイント "paint" (Okada, 2004), and occasionally to the short monophthong Jap/e/ as in /aɾeNdʒi/ アレンジ "arrange," and there are some loanwords that are read in the Japanese Romanization way as in /kanadiaN/ カナディアン "Canadian" in which the second "a" is adapted to Jap/a/, rather than the sound-based adaptation (Suzuki, 2008). Japanese ESL learners seem strongly interfered by these complex loanword adaptation processes.

In the post-test, the big difference is that the participants' vowels were much more consistently target-like than in the pre-test. This implies that they learned the idea of discreteness of phoneme in English. Above all, most of them pronounced /ei/ in "Canadian" and "arrange" correctly. This indicates that their errors regarding /ei/ are not triggered by L1 phonological constraints, but from loanword interference and a lack of knowledge of the target sound, just like the /s/ and /ʃ/ distinction. However, even in the post-test, there were still a few errors observed. A few of the participants, especially P11, P23, and P24, who had been exposed in native English relatively long among the participants, were probably tempted to add "stereotypical English flavours" and did not seem comfortable with the vowel [ei] that exists in their familiar Japanese. In other words, this is considered to be the aforementioned "avoidance" of the sounds that are similar to L1. These participants are considered to be one step ahead in the U-shaped learning (Gass & Selinker, 2009). The errors from the other participants may be just a slip of tongue.

Then, I will look at Eng/ou/. Here, some of the productions considered as errors could be within the acceptable range of the phoneme Eng/ou/. However, the productions that were obviously different from the typical Eng/ou/ in vowel length or diphthongization were counted as errors. The purpose of this criterion is that Japanese has some distinctions in the mid-back region, such as /o/, /oː/ and /oɯ/, as mentioned in 4.6. Just as the single phoneme Eng/ei/ was pronounced in various ways, Eng/ou/ also can be pronounced in various ways influenced by those Japanese distinctions and other factors.

**Table 4.10**
***Pre-test for /ou/***

| | **Target: /ou/** 4 occasions | | | | |
|---|---|---|---|---|---|
| | Toronto | don't | clothes | hope | Error |
| P1 | o(ʔ) | | ɑ | ɔ | 3 |
| P2 | oɹ | o | | ɔ | 3 |
| P3 | o(ʔ) | | | ɔ | 2 |
| P4 | | o | | | 1 |
| P5 | o(ʔ) | o | | o | 3 |
| P6 | o(ʔ) | o | o | o | 4 |
| P7 | | o | | | 1 |
| P8 | oɹn | o | ɔɹ | | 3 |
| P9 | | | | | 0 |
| P10 | | o | | | 1 |
| P11 | No V | ɑ | ɔ | | 3 |
| P12 | o(ʔ) | o | | ɔ | 3 |
| P13 | o | | | | 1 |
| P14 | | o | | | 1 |
| P15 | | o | | | 1 |
| P16 | o(ʔ) | o | | | 2 |
| P17 | o(ʔ) | o | | | 2 |
| P18 | o(ʔ) | o | | | 2 |
| P19 | o(ʔ) | o | | | 2 |
| P20 | o(ʔ) | o | | o | 3 |
| P21 | o(ʔ) | o | | | 2 |
| P22 | o(ʔ) | o | | | 2 |
| P23 | | o | | ɔ | 2 |
| P24 | No V | o | | | 2 |
| P25 | oɹ | o | | o | 3 |
| P26 | No V | o | | o | 3 |
| P27 | o(ʔ) | o | | | 2 |
| Sum | 20 | 23 | 4 | 10 | 57 |

**Table 4.11**
***Post-test for /ou/***

| | **Target: /ou/** 4 occasions | | | | |
|---|---|---|---|---|---|
| | Toronto | don't | clothes | hope | Error |
| P1 | | | ɑ | | 1 |
| P2 | | | | | 0 |
| P3 | | | | | 0 |
| P4 | | | | | 0 |
| P5 | | | | o | 1 |
| P6 | | ɑ | ɑ | ɑ | 3 |
| P7 | | | | | 0 |
| P8 | oɹ | ɚ | ɔɹ | **oɹ** | **4** |
| P9 | | | | | 0 |
| P10 | | | | | 0 |
| P11 | | | | | 0 |
| P12 | | o | | | 1 |
| P13 | | | | | 0 |
| P14 | | | | | 0 |
| P15 | | | | | 0 |
| P16 | | | | | 0 |
| P17 | | | | | 0 |
| P18 | o(ʔ) | | | | 1 |
| P19 | | | | | 0 |
| P20 | | | **oɹ** | | 1 |
| P21 | | o | | | 1 |
| P22 | | | | | 0 |
| P23 | | | | ɔ | 1 |
| P24 | | | | | 0 |
| P25 | | | | | 0 |
| P26 | | | | | 0 |
| P27 | | | | | 0 |
| Sum | 2 | 4 | 4 | 4 | 14 |

*Note.* No V=no vowel, or the vowel was not pronounced.
The bold symbols indicate errors that were not committed in pre-test.
The bold numbers of error are larger than in pre-test.

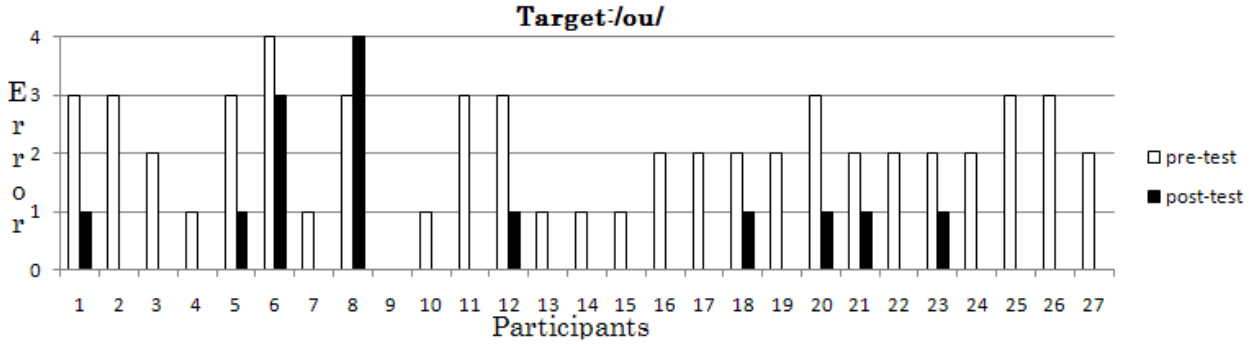


*Figure 4.8.* **The number of errors for the /ou/-stimuli in pre-test and post-test**

Between the pre-test and the post-test, the average number of errors by each participant dropped from 2.11 to 0.52. The change from the pre-test to the post-test is notable. In the pre-test, the qualities of the participants' vowels varied in a wider range than their consonants, just like /ei/. As well, there is a tendency in the error patterns. The vowel /ou/ in "Toronto" was pronounced as short [o] followed by the glottal stop [ʔ] by many participants. Japanese short vowels before pause are typically followed by [ʔ] (Vance, 2008), so it can be considered as interference from the loanword /toɾoNto/ ([toɾonto(ʔ)]) "Toronto," which finishes in the short vowel /o/, instead of /oː/ or /oɯ/. The vowel /ou/ in "don't" was also often too short to be considered correct. It is likely because in high school classroom in Japan, "don't" is predominantly pronounced as [dont(o)] with a short monophthong [o]. As well, there is a loanword /doNmai/ ドンマイ from "don't mind" in which Eng/ou/ is adapted to the short monophthong Jap/o/. This loanword adaptation may have affected the participants' productions. The vowel /ou/ in "hope" was pronounced as the pure monophthongs [oː] or [o] by some participants. It can be also considered as interference from the loanword /hoːpɯ/ ホープ "hope" in which Eng/ou/ is adapted to the long monophthong Jap/oː/. Even if they were interfered by this loanword, their [oː] was not as long as typical Japanese long vowels. It is possible that the participants tried to adopt the

L2 stress-timed rhythm, rather than the L1 mora-timed rhythm consciously or unconsciously, which made the vowel length distinction in their IL less clear, just as they did with [e:] for the word "paper." The vowel /ou/ in "clothes" was pronounced correctly by most of the participants probably because this word is not a part of the core Japanese vocabulary. Again, the participants' strong loanword interference can be observed.

Interestingly, some of the participants rhotacized /ou/ and others pronounced it as a more open vowel, like [ɔ], both of which usually do not occur in Japanese. These participants may have wanted to avoid L1-like sounds and to use more foreign sounds. In addition, P11, P24, and P26 omitted /ou/ in "Toronto." Since closed syllables occur more often in English than in Japanese, some participants may have believed that using closed syllables made their pronunciation more target-like. As a result, they pronounced "Toronto" as "Torontor" or "Toront" finishing in a closed syllable. Moreover, their rhotacization of /ou/ and omission of /ou/ in "Toronto" imply that they did not rely on the orthographical information in English. Japanese ESL learners may have to be taught that "Toronto" should be spelled as "Torontor" or "Toront" if they rhotacize or omit the final /ou/.

In the post-test, the participants' vowel qualities became much more consistently target-like again. It is considered because they learned the concept of the phonological categorization. However, a few of them still seemed tempted to add "stereotypical English flavours." This time, P6, P8, and P25 showed this tendency, although P25 tried to get rid of stereotypical English flavours by persuading herself hard before the post-test. P6 had been taught English by a native English-speaking teacher in Japan, so she had been exposed in native English for a relatively long time among the participants. Therefore, it is possible that P6 is one step ahead in the U-shaped learning. P25 also had spent a relatively long time among the participants, 8.5 months, in Canada. P25 was also possibly in the process

of U-shaped learning. However, other participants who had been exposed in native English for a relatively long time, such as P18, P24, and P26, got rid of stereotypical English flavours in the post-test. Therefore, it is not clear whether length of exposure in native English and difficulty in getting rid of stereotypical English flavours are related. As for P8, she rhotacized all the occasions of /ou/ in the post-test. In fact, she added [ɹ] to most of the words in the passage; for example, "I am" and "Canada" were pronounced as [ai aɹm] and [kʲænɑdəɹ] respectively although in the lesson, I explained that in English [ɹ] is not "spice of an English flavour," but a phoneme just like [k] or [t]. Not only P8, but several other participants did over-rhotacization to some extent. In fact, recall that all the participants misconceived that the word "idea" must be pronounced as [aidiɹ], not [aidiə], as mentioned in 4.10.2.c. Over-rhotacization seemed quite difficult to eliminate. However, over-rhotacization is also considered as a progress in U-shaped learning. In fact, they had already acquired [ɹ] to some degree, and technically, what they need to learn is the simple rule that in many dialects of North American English, when "r" appears in the spelling, [ɹ] is supposed to be pronounced, and when "r" is absent, [ɹ] is not supposed to be pronounced.

To sum up, their errors regarding Eng/ei/ and Eng/ou/ also showed that the participants lacked phonological and orthographic awareness.

### 4.13 Results of /θ/

As a control variable, I also examined three occasions of /θ/ before high front vowels, namely "think," "thing," and "healthy." In this study, I did not require the participants to learn /θ/ because this sound does not exist in Japanese as a phoneme and will require some articulatory training. Therefore, I did not mention this phoneme at all during the lesson. In fact, the difference between the pre-test and post-test was not significant ($p > 0.5$). This

indicates that their improvement of /s/, /ʃ/, /z/, /ei/, and /ou/ are because of the 40-minute lesson, not simply because time went by. More importantly, there are some differences in their consonant substitution patterns between the pre-test and post-test. As shown in Table 4.12, P2, P11, P13, P16, and P21 were biased toward [s] (or against [ʃ]), while P5 and P7 were biased toward [ʃ] in the pre-test. After the lesson, they became unbiased, or possibly became biased toward the other side, which affected the substitution for /θ/.

**Table 4.12**
*Pre-test for /θ/*

| Target:/θ/ 3 occasions | | | | |
|---|---|---|---|---|
| | think | thing | healthy | Error |
| P1 | s | s | | 2 |
| P2 | s | | s | 2 |
| P3 | s-θ | | s | 2 |
| P4 | | | | 0 |
| P5 | | | ʃ | 1 |
| P6 | | | s | 1 |
| P7 | | s | ʃ | 2 |
| P8 | s | s | s | 3 |
| P9 | | | | 0 |
| P10 | | | | 0 |
| P11 | | s-θ | s | 2 |
| P12 | s | s | s | 3 |
| P13 | s | s | s | 3 |
| P14 | s | s | s | 3 |
| P15 | s | | s | 2 |
| P16 | s | | s | 2 |
| P17 | s | s | s | 3 |
| P18 | s | | s | 2 |
| P19 | s | | s | 2 |
| P20 | | | | 0 |
| P21 | s | s | s | 3 |
| P22 | s | s | s | 3 |
| P23 | | t | | 1 |
| P24 | | | | 0 |
| P25 | | | | 0 |
| P26 | | | s | 1 |
| P27 | | | | 0 |
| Sum | 14 | 11 | 18 | 43 |

**Table 4.13**
*Post-test for /θ/*

| Target: /θ/ 3 occasions | | | | |
|---|---|---|---|---|
| | think | thing | healthy | Error |
| P1 | s | s | | 2 |
| P2 | ʃ | | ʃ | 2 |
| P3 | | s | s | 2 |
| P4 | | | | 0 |
| P5 | | | s | 1 |
| P6 | | | | 0 |
| P7 | | | s | 1 |
| P8 | s | s | s | 3 |
| P9 | | | | 0 |
| P10 | | | **s** | **1** |
| P11 | **ʃ** | s | | 2 |
| P12 | s | s | s | 3 |
| P13 | s | s | ʃ | 3 |
| P14 | s | s | s | 3 |
| P15 | s | | s | 2 |
| P16 | ʃ | **s** | ʃ | **3** |
| P17 | s | s | s | 3 |
| P18 | s | **s** | s | **3** |
| P19 | s | | ʃ | 2 |
| P20 | | | | 0 |
| P21 | ʃ | ʃ | ʃ | 3 |
| P22 | s | s | s | 3 |
| P23 | | t | | 1 |
| P24 | | | | 0 |
| P25 | | | | 0 |
| P26 | | | s | 1 |
| P27 | | | | 0 |
| Sum | 14 | **13** | 17 | **44** |

## 4.14 Summary

When Japanese ESL learners in my study mispronounced /s/ and /ʃ/ before high front vowels, they simply did not know what the target sounds are. Once the Japanese ESL learners were taught the basic symbol-sound correspondence rules, namely "s"-/s/, and "sh"-/ʃ/, they were no longer totally confused between the two sounds. No articulation training was necessary. The same was true for /z/, /ei/, and /ou/, except Japanese ESL learners might need articulatory practice for the pure voiced fricative [z]. In other words, it is enough simply to explain English spelling rules. However, Japanese ESL learners also need to be reminded that Japanese has the counterparts of these English sounds because some participants had been convinced that /s/ and /ʃ/ were their weakness. The source of their pronunciation errors regarding /s/, /ʃ/, and /z/ before high front vowels, and the diphthongs /ei/ and /ou/ was a lack of phonological awareness in English and the knowledge of the symbol-sound correspondence rules. Therefore, they often could not divide English words into phonemes or other phonological units. In other words, to those learners, English words often did not consist of phonemes in both production and perception. Moreover, Japanese ESL learners in my study seemed aware that they lacked the knowledge of the English symbol-sound correspondence rules and the idea of discreteness of phoneme in English. In fact, the participants might benefit from the lesson on the English phonological categorization and symbol-sound correspondence rules. On the other hand, the participants seemed flexible about phonetic characteristics. What is important is that the 27 participants were all from different high schools in various regions in Japan, but none of them had been formally taught the very basic symbol-sound correspondence rules of English during the six-year English curriculum. This fact needs to be seriously considered when pronunciation materials for Japanese ESL/EFL learners are designed.

# Chapter Five

# SUMMARY AND CONCLUSION

## 5.1 Summary

It is a general belief in Japan that the English /s/ and /ʃ/ before high front vowels are problematic for Japanese ESL/EFL learners. English pronunciation textbooks in Japan often say that Japanese has [ʃi] but does not have [si]. Since some of these English pronunciation textbooks tend to greatly look down upon the Japanese sound system, some distinctions that are widely believed missing in Japanese, including the [s] and [ʃ] distinction, may actually exist in Japanese. In Japanese, there seem to be no minimal pairs of [si] and [ʃi] in core lexical items, whereas in peripheral lexical items, such as some cultural dialects and names, [si] behaves like a distinctive syllable, meaning that /si/ and /ʃi/ are contrastive.

According to the experiment with the Japanese contexts, 93 monolingual Japanese speakers between the ages of 17 and 89 all distinguished [si] and [ʃi] in both perception and production. In contrast, according to the experiment with the English contexts, in the pre-test, 23 out of the 27 Japanese ESL learners made errors regarding /s/ and /ʃ/ when reading an English passage. Interestingly, their errors were caused by a lack of phonological and orthographic awareness in English, not by an inability to articulate. In addition, the participants were not able to utilize the IPA. Once they were reminded that Japanese has the /s/ and /ʃ/ distinction and were taught some basic symbol-sound correspondence rules, such as the rules that “s” corresponds to /s/ and “sh” correspond to /ʃ/, most of the participants immediately corrected their pronunciation without articulatory practice. In the post-test, 15 out of the 27 participants did not make even a single mistake among both

[s]-stimuli and [ʃ]-stimuli. The number of the total errors regarding [s] and [ʃ] by all the participants dropped from 106 to 26 between the pre-test and the post-test. This change is significant. The same logic applies to /ei/ and /ou/.

In terms of phonetic aspects, in the Japanese contexts, about 50% of Japanese speakers (both monolingual Japanese speakers and Japanese learners or English) sometimes pronounced /si/ as [swi] or the like. However, in the English contexts, this on-glide-insertion was not observed. As for /ʃ/, since acoustically the Japanese /ʃ/ is higher in frequency than its English counterpart, post-pausal /ʃ/ in the Japanese contexts, especially produced by young females, can sometimes be categorized as the English /s/. However, in the English contexts, once the Japanese ESL learners are aware of the phonemic concept of /s/ and /ʃ/, [s]-like /ʃ/ seldom appears. As for /s/, although the Japanese /s/ is generally said to be pronounced like [θ], in the English contexts, it does not completely sound like the English /θ/. While the Japanese ESL learners in my study often lack phonological awareness, they seem to be sensitive to phonetic qualities. This could be an unexpected benefit of lacking phonological awareness. Therefore, I argue that the cross-linguistic phonetic differences regarding /s/ and /ʃ/ do not need to be taught to Japanese ESL learners unless learners are eager for native-like pronunciation or until learners feel comfortable with the phonemic difference and the symbol-sound correspondence rules regarding /s/ and /ʃ/.

This study found that Japanese speakers are able to distinguish between /s/ and /ʃ/ before high front vowels, implying that authors of English pronunciation textbooks in Japan may have to reconsider whether they should write that the /s/ and /ʃ/ distinction is one of the weaknesses of Japanese ESL/EFL learners.

## 5.2 Pedagogical Implications

Based on my study, there is a possibility that Japanese ESL/EFL learners often have not yet been taught the basics of the English phoneme inventory and the symbol-sound correspondence rules. Japanese ESL learners in my study cannot utilize IPA, either. According to the experiments mentioned in Chapters Three and Four, basic knowledge of the English phoneme inventory and the symbol-sound correspondence rules helps Japanese ESL learners improve their pronunciation regarding /s/, /ʃ/, /ei/, and /ou/ without articulatory practice. Based on the lesson I gave during the experiment mentioned in Chapter Four, the following is the process that I think would help Japanese ESL/EFL learners understand the basics:

1. Remind Japanese ESL/EFL learners that Japanese-speaking children learn the Japanese mora inventory with each *kana* symbol that corresponds to each mora (the Japanese phonological unit) in elementary school.
2. Explain that English-speaking children take the same types of lessons in elementary school as well.
3. Show the English phoneme inventory with the chart in Figure 4.3 or equivalents.
4. Encourage Japanese ESL/EFL learners not to use other sounds than those in the English phoneme inventory. This advice would greatly narrow down Japanese ESL/EFL learners' choices of English sounds.
5. Explain that there are basic symbol-sound correspondence rules that are completely different from the Japanese Romanization rules. For example, the letter "s" usually corresponds to /s/, the letters "sh" usually correspond to /ʃ/, the letter "a" usually corresponds to either /ei/ or /æ/, the letter "o" corresponds to either /ou/ or /ɑ/, and so forth.

6. Teach that phonemes change lexical meanings by showing phonemes that are familiar to Japanese ESL learners. For example, /p/ and /b/ are different phonemes, as in "pan" and "ban." Likewise, /s/ and /ʃ/ are two different phonemes, which change lexical meanings. It is not the story that [ʃi] sounds more Japanese-like and [si] sounds more English-like.
7. Encourage Japanese ESL/EFL learners to utilize as many Japanese sounds (including allophones) that are similar to particular English sounds as possible, and minimize effort of articulatory practice.[45] Since Japanese speakers can distinguish between [si] and [ʃi], it seems unnecessary to teach articulation of these sounds.
8. Note that the learner's purpose at beginner level is to produce the target phoneme within the acceptable range. Native-like production may be the second step. Even if a learner's terminal goal is native-like production, the learner has to achieve intelligibility as an intermediate goal.

English pronunciation textbooks mainly focus on articulation. Therefore, based on Fraser (in press), it would be better to focus on building learners' *understanding* of the basic phonological systems of English. In addition, based on Archibald (2005), it would be efficient to get rid of as many unnecessary articulatory explanations as possible and to economically utilize L1 sounds.

## 5.3 Limitations

- In both Experiment A and B, the judges listened to the recordings together and discussed. Therefore, the judges may have influenced each other. I did not calculate inter-rater reliability. On the other hand, discussions enabled us to notice some

[45] The basis of this idea is from Archibald's (2005) redeployment of L1 phonological knowledge.

details: for example, /ou/ in “don’t” pronounced by many participants did not sound right because of the vowel length, rather than the vowel quality.

- In Experiment B, the post-test was done right after the lesson and there were no delay tests. Therefore, there is no guarantee that the participants retained the new knowledge afterwards. Despite this limitation, this study found that the participants were able to distinguish between /s/ and /ʃ/ before high front vowels in production.
- In Experiment B, as shown in Appendix B: Post-test, pronunciation tips are written on the sheet the participants read in the post-test. Especially, “ch” in “machine” was indicated with the letters “sh,” and /ei/ and /ou/ were indicated with the Japanese orthography. Although I taught the participants some symbol-sound correspondence rules, specifically so called “Soft-CH” says /ʃ/, “Long-A” says /ei/, and “Long-O” says /ou/, it is possible that the participants simply read the pronunciation tips, instead of the English orthography. However, this study still found that the participants were able to produce /ʃ/, /ei/, and /ou/. As for the letters “s” and “sh,” the pronunciation tips on the sheet were just reminders that the words had “s” or “sh.” Therefore, this study shows that the participants learned the rules that the letter “s” says /s/ and the letters “sh” say /ʃ/. However, there is no guarantee that the participants can utilize this new knowledge about “s” and “sh” without the reminders.
- What was taught in the lesson in Experiment B was deploying L1 phonological counterparts, but language-specific phonetic characteristics were ignored. In other words, the participants have learned to pronounce the target phonemes, but they still retained foreign accent. It may be debatable if teachers should teach native-like pronunciation or just intelligible pronunciation. From the phonetic side, the

difference between English and Japanese in phoneme boundary between /s/ and /ʃ/, and the phonetic difference between Jap/ei/ or Jap/ow/ and Eng/ei/ or Eng/ou/ can be a big issue. From the applied side, on the other hand, as long as L2 learners' accent does not impede communication, foreign accent does not necessarily need to be a worry. In case of my participants, they had been taught only phonetic details to some extent prior to the experiment. Therefore, they lacked phonological awareness and were tempted to use non-L1-like sounds, which were actually not L2-like, either. In their case, I propose they can forget native-like pronunciation or phonetic details at this moment, but instead learn to make use of L1 sounds as much as possible and focus on the phonological system. Once they build some foundations, they may have to practice articulation to aim for native-like pronunciation if they want to.

- An interview was done in Experiment B. I did not interview one by one, but I asked a few participants in the room at the same time. Therefore, the participants' answers might have been influenced by others.
- Japanese has many dialectal variations. For example, some northern dialects do not distinguish between [ʃi] and [sɯ] (Kawamoto, 1971). This present study does not cover speakers of these dialects.

## 5.4 Further Research

English pronunciation teaching methods in Japan must change. This present study found a simple fact that orthographic knowledge helps Japanese ESL learners correct their pronunciation errors of particular sounds. The question is why such a simple teaching method has not been commonly performed in Japan. Are there any factors that prevent teachers from teaching symbol-sound correspondence rules? To what extent do teachers themselves explicitly know English and Japanese pronunciation? These questions are left

for further research. As mentioned in sections 2.9 and 2.10, Japanese teachers seem prone to focus more on weaknesses than on strengths, and there seems to be a belief that the Japanese language is inferior to English (or any other language) in Japan. These social characteristics may be one of the key factors that delay the development of the pronunciation teaching methods. It is worth examining.

More specifically, how to teach English spelling rules is one of the most important issues. In this present study, I used the chart shown in Figure 4.3, but it needs to be improved. Phonics is one of the candidate methods. In Japan, some teachers seem to strongly insist on the necessity of phonics, while others are strongly critical of phonics. Another issue is that understandings of phonics may differ from teacher to teacher in Japan. According to Wardhaugh (1970), although Bloomfield and Fries were critical of phonics, they stressed symbol-sound correspondence, which is at the heart of phonics. Therefore, there is no doubt that teaching symbol-sound correspondence to Japanese ESL/EFL learners is crucial. How about IPA? The issue is the understanding of IPA by Japanese teachers of English. More importantly, Kurimoto (2009), an instructor at Kandagaigo University, stated in her presentation that IPA used in English-Japanese dictionaries needs to be improved. Makino (2009) also stressed the problems of IPA used in Japan. Why has IPA used in Japan not been improved? Are there any social issues involved? What prevents the improvement of the teaching methods may be as important as how to improve them.

To improve the phonological aspects, phonologists, especially those who specialize in second language phonology, would have to be involved much more. As mentioned in 5.2.7, it would be economical to utilize as many Japanese sounds that are similar to particular English sounds as possible in order to minimize effort of articulatory practice. What

phonologists need to research is to examine what Japanese sounds can be utilized in English and what articulatory practice that is generally done is unnecessary. This present study found that the /s/ and /ʃ/ distinction before high front vowels, the tense /ei/ and the lax /ɛ/ distinction, and the tense vowel /ou/ do not require articulatory practice although learners still retain foreign accent. Researchers need to find many more such cases.

In terms of the qualities of Jap/s/ and Jap/ʃ/, as mentioned in Chapter Three, a few Japanese speakers tend to pronounce [θ]-like /s/, and several Japanese speakers, especially females in their teens, 20's and 30's tend to pronounce [s]-like /ʃ/ especially in a post-pausal position. It would be needed to examine how native speakers of English perceive those sounds by performing perception tests.

## 5.5 Conclusion

The primary goal of this study is to examine whether the /s/ and /ʃ/ distinction before high front vowels is difficult for Japanese ESL/EFL learners. The following are the answers for the research questions mentioned in 1.2.

1. What phonological status does [si] have in Japanese?

   In Japanese, [si] and [ʃi] seem to be recognized as clearly different sets of sounds by Japanese speakers. In peripheral classes of lexicon (e.g., foreign names, jargons, nicknames, names of characters in comic books, and computer games) I argue that [si] and [ʃi] are contrastive and some minimal pairs exist. In relatively central lexical items, I argue that [si] is still considered to be a foreignism. On the other hand, [si] and [sʷi] are not contrastive.

2. Can monolingual Japanese speakers produce both [si] and [ʃi]?

   Yes, they can. Monolingual Japanese speakers in or around Tôkyô in my study

distinguished between [si] and [ʃi] in both reading tasks and mimicking regardless of their age.

3. Can monolingual Japanese speakers distinguish between the [si] and [ʃi] syllables through listening?

   Yes, they can. Monolingual Japanese speakers in or around Tôkyô in my study perceived the difference between [si] and [ʃi] produced by a native speaker of North American English regardless of their age. Moreover, monolingual Japanese speakers perceived the English [ʃi] more accurately than North American English speakers.

4. If the Japanese speakers can produce and perceive the difference between [si] and [ʃi], what then would be the sources of Japanese speakers' challenges in mastering the distinction in their English production?

   The /s/ and /ʃ/ confusion by the Japanese ESL learners in my study is caused by a lack of phonological and orthographic awareness in English, not by an inability to articulate. In other words, they do not know when to pronounce /s/ and /ʃ/. Once they are reminded that Japanese has the /s/ and /ʃ/ distinction and are taught some basic symbol-sound correspondence rules, such as the rules that "s" says /s/ and "sh" say /ʃ/, they can correct their pronunciation without articulatory practice. Therefore, it is enough simply to explain English spelling rules. In other words, Japanese ESL/EFL learners' challenges are to learn the basic English phonological and orthographic rules without being distracted by loanword interference and personal preference.

To conclude, English pronunciation teaching methods in Japan must change. Ironically,

some problems are not due to L1 constraints, but seem to be made up by misleading teaching. Specifically, pronunciation textbooks have to focus more on the basic phonological rules, or *when* to pronounce each phoneme in each word, not *how* to pronounce it. The textbooks also need to clearly explain the basic English symbol-sound correspondence rules. Since authors of pronunciation textbooks are not necessarily specializing in phonology, their understanding of the basic phonological rules of both English and Japanese may not be always accurate. To solve this problem, phonologists need to be involved more in the English education to help these authors and teachers of English understand the basics.

## Bibliography

Adachi, T., & Yamada, T. (1998). *Êgo Listening Kagakuteki Jôtatsuhô*. Tôkyô, Japan: Kôdansha.

Agency for Cultural Affair. (2002). *Kokugo-hyôki no kijun*. Retrieved March 30, 2009, from http://www.bunka.go.jp/kokugo/main.asp?fl=list&id=1000003933&clc=1000000068

Akahane-Yamada, R., Kubo, R., Nishi, K., Strange, W., & Trent, S. A. (2001). Effects of consonantal context on perceptual assimilation of American English vowels by Japanese listeners. *Acoustic Society of America*, *109*(4). 1691-1704.

Akamatsu, T. (2000). *Japanese phonology*. LINCOM Studies in Asian Linguistics 38.

Akinaga, K., & Kindaichi, H. (Eds.). (2008). *Shinmêkai Nihongo accent jiten*. Tôkyô, Japan: Sansêdô.

Akiyama, K. (2009). *Sensê seminar: Hatsuon hassê*. NHK CTI Nihongo Centre.

Archibald, J. & O'Grady, W. (2004). *Contemporary linguistics analysis*. Ontario: Pearson Education Canada Inc.

Archibald, J. (2005). Second language phonology as redeployment of phonological knowledge. *Canadian Journal of Linguistics, 50*, 285-314.

Arimoto, J. (1993). Tactics of error correction in English pronunciation: Consonants. *Sonoda Women's College Studies. 27*, 75-90.

Arimoto, J. (2002). Teaching materials in English pronunciation. *Kansai University of International Studies*, *3*.

Avery, P. & Ehrlich, S. (2003). *Teaching American English pronunciation*. New York: Oxford University Press.

Bauer, L. & Trudgill, P. (1998). *Language myths*. London: Penguin Group.

Beckman, M. E., Edwards, J., & Li, F. (2009). Contrast and covert contrast: The phonetic

development of voiceless sibilant fricatives in English and Japanese toddlers. *Journals of Phonetics, 37*, 111-124.

Berman, J., Lambacher, S., Martens, W. & Nelson, B. (2001). Identification of English voiceless fricatives by Japanese listeners: The influence of vowel context on sensitivity and response bias. *Acoustical Science and Technology, 22*(5), 334-343.

Brown, H. D. 2007. *Principles of language learning and teaching*. New York: Pearson Education.

Carroll, J. M. (2004). Letter knowledge precipitates phoneme segmentation, but not phoneme invariance. *Journal of Research in Reading*, *27*(3), 212-225.

Davenport, M. & Hannahs, S. J. (2005). *Introducing phonetics & phonology.* London: Hodder Education.

Duran, J. (2005). Tense/Lax, the vowel system of English and phonological theory. *Headhood, Elements, Specification and Contrastivity: Phonological Papers in Honor of John Anderson, 77*.

Eckman, F. R. (2008). Typological markedness and second language phonology. In J. G. Hansen Edwards & M. L. Zampini (Eds.), *Phonology and Second Language Acquisition* (pp. 95-116). Philadelphia: John Benjamins.

Elbert, S. H. (1970). *Spoken Hawaiian.* Honolulu: University of Hawai'i Press.

Fayol, M. & Pacton, S. (2004). Learning to spell in a deep orthography: The case of French. In R. Berman & T. Gillis (Eds.), *Trends in language acquisition research*. Dordrecht: Kluwer.

Fraser, H. (in press). Pronunciation as categorization: The role of contrast in teaching English /r/ AND /l/. *Australian Review of Applied Linguistics*.

Fukasawa, T. (2007). *A guide to perfect English pronunciation*. Tôkyô, Japan: ALC.

Fukushima, E., Matsuzawa, K., &. O'Connor, C. (2008). *Êgomimi boitore*. Tôkyô, Japan: ASCII MEDIA WORKS.

Fukushima, E. (2009). *Kirêni hanasutameno hassê hatsuon training*. Tôkyô, Japan: Bere-shuppan.

Gass, S. M. & Selinker L. (2009). *Second language acquisition*. New York: Routledge.

Goad, H. & Kang, H. (2003). Word-final syllabification in L2 acquisition with emphasis on Korean learners of English. *Generative approaches to second language acquisition conference*. pp. 122-129. Somerville, MA: Cascadia Proceeding Project.

Goble, D. (2002). Loanword-induced interference in Japanese students' foreign language acquisition: Developing student awareness through experiential learning. *Journal of Chikushi Jyogakuen Junior College*, *37*, 55-77.

Godin, B. (2002). *The sound & rhythm of English*. Tôkyô, Japan: Shinnihonkyôikutosho.

Gore, M. (1998). Phonetic analysis of Japanese university students' pronunciation of English vowels. *Bulletin of the Faculty of Education,Kagoshima University. Cultural Social Sciences*, *49*, 117-125.

Gore, M. (2004). Phonological and phonetic views of vowel sequence timing : A cross-linguistic review. *Bulletin of the Faculty of Education, Kagoshima University. Studies in education*, *55*, 83-99.

Grenon, I. (2005). The status of the sound [z] in Japanese. *Revue Langues et Linguistique*, *31*, 64-90.

Grenon, I. (2008). The acquisition of English sound [dz] by native Japanese speakers. *A Perceptual Study*. Saarbrücken, Germany: VDM Verlag.

Haraga, M. (2009). Kokodakewa naose! Nihonjin hatsuon. *AERA English 45*. Tôkyô, Japan: Asahishinbunsha.

Haraguchi, Y. (2003). The acquisition of aspiration of voiceless stops and intonation patterns of English learners: Pilot study. *Proceeding of the 8th conference of Pan-Pacific Association of Applied Linguistics*, 83-91.

Hirata, Y. (2004). Effects of speaking rate on the vowel length distinction in Japanese. *Journal of Phonetics, 32*(4), 565-589.

Hirose, K., Minematsu, N., Okabe, K. & Shyu, H. (2005). Experimental study on objective intelligibility of Japanese English to American listeners. *IEIC Technical Report*, *104*(630), 31-36.

Hisamatsu, T. & Nitta, Y. (1994). An efficient treatment of Japanese verb inflection for morphological analysis. In *COLING 94*.

Hockett, C. F. (1960). The origin of speech. *Scientific American, 203*, 88-96.

Horiguchi, M. (1997). An example of introducing the English phonetics: By citing Japanese loanwords from English. *The Journal of Science of Culture and Humanities and Nature of Shin-ai Women's Jr. College, 37*. 39-47.

Huang, L. (2003). Ten pointers for enhancing learners' motivation. *Business communication quarterly, 66*(4), 88–95.

Ikegaya, Y. (2008). *Kowai kurai tsûjiru Katakana Êgono hôsoku*. Tôkyô, Japan: Kôdansha.

Inoue, H., Maruya, S., Ôno, S., & Ôoka, M. (1990). *Nihongo sôdan 1*. Tôkyô, Japan: Asahishinbunsha.

Inozuka, E., Inozuka, H., & Machida, K. (Ed.). (2009). *Nihongo onsêgakuno shikumi*. Tôkyô, Japan: Kenkyûsha.

Irie, S. (2004). The phonetic structure of adjective stems in modern Japanese: An analysis and study of phonemic frequency. *Bulletin of Center for Japanese Language Dôshisha University.*

Itô, J. & Mester, A. (1995). The core-peripheral structure of the lexicon and constraints on reranking. *University of Massachusetts Occasional Papers*, *39*, 181-209.

Itô, J. & Mester, A. (2003). Systemic markedness and faithfulness. *Chicago Linguistic Society*, *39*(1), 556-689.

Jolly Learning Ltd. (2008). *Jolly Phonics*. Retrieved June 20, 2010, from http://www.jollylearning.co.uk/jp.htm

Jones, M. L. (1996). Phonics in ESL literacy instruction: Functional or not?. *Proceedings of the 1996 World Conference on Literacy.*

Kager, R. (2006). *Optimality heory*, New York: Cambridge University Press.

Kawamoto, E. (1971). Toyama-ken Shôgawa-ryûikini okeru Zûzû-benbo bunputo sono kaishaku. *Kanazawa Daigaku Gogaku Bungaku Kenkyû 2*, 1-10.

Kim, C. W. (1968). The vowel system of Korean. *Language*, *44*(3), 516-527.

Kindaichi, H. (1989). *Nihongo.* Tôkyô, Japan: Iwanami Shoten.

Kindaichi, H., & NHK Hôsôbunka Kenkyûjo (Ed.). (1997). Kyôtsûgono hatsuonto accent. *Nihongo Hatsuon Accent Jiten*. Tôkyô, Japan: NHK Shuppan.

Kingston, J. (2007). The phonetics-phonology interface. In: de Lacy, P. (ed). *The Cambridge Handbook of Phonology*, 401-435. Cambridge: Cambridge University Press.

Kobayashi, M. (2007). On the primordial vowel system of Japanese: A semantic approach. *International Buddhist University bulletin, 45*, 379-411.

Koizumi, T. (2003). *An introduction to phonetics*. Tôkyô, Japan: Daigakushorin.

Ladefoged, P. (2006). *A course in phonetics*. Boston: Thompson Wadsworth.

Lee, S. (2009). Homophony avoidance effect: Phonetic studies on contrastiveness in Korean vowel length. *Harvard Studies in Korean Linguistics XIII.*

Lin, H. (2005). Understanding problems in learning Mandarin consonants by monolingual

speakers of English. *Journal of Canadian Teachers of Chinese as a Second Language, 1*(15).

Lin, H. (2008, April 18-20). Interlanguage rhythm in the English production of Mandarin speakers. *Proceedings of the 8th Phonetic Conference of China and the International Symposium on Phonetic Frontiers*. Beijing.

Lovins, J. B. (1975). *Loanwords and the phonological structure of Japanese*. Indiana University Linguistics Club.

Maddieson, I. (1984). *Patterns of sounds*. New York: Cambridge Press.

Magnuson, T. J. (2008). *What /r/ sounds like in Kansai Japanese: A phonetic investigation of liquid variation in unscripted discourse*. MA Thesis. University of Victoria.

Makino, T. (2008a). English pronunciation teaching and phonetic study. *Proceedings of the 162th Meeting of Osaka Linguistics Society.* Ôsaka, Japan.

Makino, T. (2008b). Consonant substitution patterns in Japanese speakers' English. *Proceedings of II International Conference on Native and Non-native Accents of English.* Łódź, Poland.

Makino, T. (2009). Do we need British variants in the pronunciation entries of learner's English-Japanese dictionaries?. *manuscript*.

Mase, Y. & Nakato, Y. (1998). Orthographic problems of loanword in Japanese education: With reference to Korean learners of Japanese. *Ferris Studies*, *33*, 111a-85a.

Masuda, Y. (2003). An experimental study on perceptual difficulties of English consonants by Japanese learners of English. *Studies in Language and Literature, 22*(2), 119-150.

Matsuzaki, H. (1992). A typification of variation in loanword sounds in dictionary items. *Tsukuba Working Paper in Linguistics, 10/11*, 43-56.

Matsuzaki, H. (1993). A quantitative study of fluctuation in Japanese loanword-phone

notation. *Journal of the Department of Japanese, Tohoku University, 3,* 83-94.

Melenca, A. (2001). Teaching connected speech rules to Japanese speakers of English so as to avoid a staccato speech rhythm. *A Thesis in the Centre for Teaching English as a Second Language*. Concordia University.

Miyauchi, T. (1993). Implementation of English word retrieved system with Katakana. *Jôhôshorigakkaikenkyûhôkoku NL* 91-17.

Muroi, A. (2005). Hatsuon shidôni okeru kyôshi no yakuwari: Ayashii hatsuon shidô no shôtai. *English teachers' magazine, 54*(10). Tôkyô, Japan: Taishûnshoten.

Nakayama, K., & Yamagushi, K. (2003). The confusion of alveolar and palato-alveolar fricatives by Japanese learners of English. *Bulletin of the Yamagata University, Cultural Science,15*(2), 207-220.

NHK Hôsôbunkakenkyûjo (Ed.). (2008). *Nihongo hatsuon accent jiten*. Tôkyô, Japan: NHK Shuppan.

NHK Shuppan (Ed.). (2005). NHK announce jissen training. Tôkyô, Japan: NHK Shuppan.

Nogita, A. (2010). Do Japanese ESL learners' pronunciation errors come from inability to articulate or misconceptions about target sounds?. *Working Papers of the Linguistics Circle of the University of Victoria 20*, 82-116.

Ohata, K. (2004). Phonological difference between Japanese and English. *Asian EFL Journal: English Language Teaching and Research Articles*, *6*(4).

Okada, S. (2004). Language variation of the hiatus /ei/ in the Corpus of Spontaneous Japanese-in the case of loanwords-. *IEICE Technical Report. Speech, 104*(148), 35-40.

Ôno, S. (1994). *Nihongoni tsuite*. Tôkyô, Japan: Iwanami-shoten.

Ôno, S. (2002). *Nihongowa ikanishite sêritsu shitaka*. Tôkyô, Japan: Chûkô-bunko.

Ôno, S. (2003). *Nihongowa dokokara kitanoka*. Tôkyô, Japan: Chûkô-bunko.

Ôno, S. (1994). *Nihongono genryûo motomete*. Tôkyô, Japan: Iwanami-shoten.

Pan, H. Y., Utsugi, A., & Yamazaki, S. (2004). An acoustic phonetic study of voiceless alveolo-palatal fricatives in Japanese, Korean and Chinese. *Journal of General Linguistics, 7*, 1-27.

Rysiewicz, J. (1990). The English palatalization rule in second language acquisition. *Volume Papers and Studies in Contrastive Linguistics, 26,* 111-122.

School Specialty Publishing (2007). *Phonics: Grade 1*. Ohio: School Specialty Publishing.

Scobbie, J. M. & Stuart-Smith, J. (2006). Quasi-phonemic contrast and the fuzzy inventory: Examples from Scottish English. *QMU Speech Science Research Centre Working Papers.*

Seliger, H. W. & Shohamy, E. (2003). *Second language research methods*. New York: Oxford University Press.

Shiba, S. (2008) *Nihongono kigen: Sono gutaiteki zentaizô*. Tôkyô, Japan: San'ichi-shobô.

Steele, J. (2002). L2 learners' modification of target language syllable structure: Prosodic licensing effects in interlanguage phonology. In A. James & J. Leather (Eds.), *New Sounds 2000: Proceedings of the fourth international symposium on the acquisition of second-language speech* (pp. 315-324). Klagenfurt, Austria: University of Klagenfurt.

Stephens, M. A. (2002). The interference of written English on the pronunciation of EFL learners in Japan. *Studies in Language and Literature*, *22*(1), 87-111.

Suzuki, S. (2008). A study of Japanized English : Its structures and thought. *Bulletin of Kokusai Junior College 23*, 1-47.

Takahashi, E. (2005). Listening errors among Japanese EFL learners, *Studies of Regional Policy 8(1),* 45-53.

Takahata, Y. (2003). *Sonokini saseru coachingjutsu.* Tôkyô, Japan: Sankaidô.

Takayama, M. (2003). Issues in Japanese historical phonology. *Journal of the Phonetic Society of Japan, 7*(1), 35-46.

Tanaka, Y. & Suarez, A. (2001). Japanese learners' attitudes toward English pronunciation. *Bulletin of Niigata Seiryo University, 1*, 99-111.

Tatsumi, I. (2009). *Êgono hatsuonga yokunaru hon*. Tôkyô, Japan: Chûkê-shuppan.

Teshigawara, M (2010, March 11). The voices of Anime. *Interdisciplinary Day: Computer Science & Linguistics*. University of Victoria.

Thayne, D. (2009). *Mireba wakarunoni nazeka kikitorenai Êgo*. Tôkyô, Japan: NTT-shuppan.

*The Speech Accent Archive*. Retrieved April 3, 2010, from http://accent.gmu.edu/browse_native.php?function=detail&languageid=33

Tsujimura, N. (2007). *An introduction to Japanese linguistics*. Oxford: Blackwell.

Tsuzuki, M. (1996). An allophonic study of Japanese vowels and consonants. *The Journal of Aichi Gakuin University. Humanities & sciences, 43*(2), 41-66.

TuZino, T. (2008). *Nihonshiki Rômaji honji*. Retrieved May 16, 2010, from http://staff.aist.go.jp/t.tuzino/doc/aikitu/web/web.html

Uekawa, K., & George, J. (2007). *Êgo nodo 50no method*. Tôkyô, Japan: Sanshûsha.

Wang, H. S. (1991). An orthography-based symbolic system for TESL, *English Teaching Forum 24*(3), 16-21.

Wardhaugh, R. (1970). Linguistics and phonics. *International Reading Association Concention*, Anaheim, Cal.

Wilson, I. (2006). Articulatory settings of French and English monolingual and bilingual speakers. *PhD dissertation*, University of British Columbia.

Vance, T. J. (2008). *The sound of Japanese*. New York: Cambridge University Press.

## Appendix A: Pre-test

以下の英文を読んで下さい。つっかえたり間違えたら言い直してもかまいません。なめらかさよりハッキリさ重視でお願いします。

I am staying in Toronto in Canada. I think Canadian people are very nice. I like to be able to see the beautiful mountains and the sea. I don't like washing machines because I have to pay with coins. One thing I like, however, is that they have machines that dry clothes.

I have a Canadian friend named Venus. We have a good relationship. Every day, Venus eats Italian food. She likes to paint cute kittens on colorful sheets of paper. She usually sits on the seat next to me in class. She likes music and arranges her CD in alphabetical order. However, she gets sick easily. She is sick right now. I hope she becomes healthy soon.

**Appendix B: Post-test**

レクチャーを思い出し、下線部の音に注意しながら、以下の英文を読んで下さい。

I am stay(エイ)ing in Toron(オウ)to in Canada. I think Ca(エイ)nadian people are very nice. I like to be (エイ)able to see the beautiful mountains and the sea. I don't(オウ) like wa(SH)shing ma(SH)chines because I have to pay(エイ) with coins. One thing I like, however, is that they(エイ) have ma(SH)chines(オウ) that dry clothes.

I have a Ca(エイ)nadian friend na(エイ)med Venus. We have a good re(エイ)lationship. Every day(エイ), Venus eats Italian food. She likes to paint(エイ) cute kittens on colorful sheets of pa(エイ)per. She usually sits on the seat next to me in class. She likes mu(Z)sic and ar(エイ)ranges her C(S)D in alphabetical order. However, she gets sick ea(Z)sily. She is sick right now. I hope she(オウ) becomes healthy soon.

## Appendix C: Interview

1. **Questions Regarding the Participants' Background**

   a. Have you ever been formally taught IPA in Japan? Can you actually utilize it?

   b. Have you ever been formally taught "phonics," "42-sounds of English," "learn to read," or an equivalent?

   c. Have you ever been formally taught English pronunciation in Japan?

2. **Quizzes**

   a. Can you identify whether the following words in each pair are the same or different in pronunciation? If different, can you identify the difference?

   - "meat" versus "meet"
   - "ear" versus "year"
   - "bone" versus "born"
   - "cars" versus "cards"
   - "branch" versus "brunch"

   b. Do you know that in English, unstressed vowels are reduced to Short U [ə]? E.g. "J<u>a</u>pan," "comm<u>e</u>nt," "terr<u>i</u>ble," "comm<u>o</u>n" and so on.

   c. For the word "idea," which do you think sounds more proper, [aɪdiə] or [aɪdiɹ]?

3. **Question for the Participants' Opinion**

Do you agree with this process?

***Katakana*-English-based underlying representation**: リード /ɾi:do/ "read"

↓

**Stereotypical L2 flavours**: "ド [do] sounds Japanese-like. ドゥ [dɯ] sounds exotic."

"The more tongue rolling (rhotacized), the more English-like."

↓

**Surface representation in IL**: [ɹi:dɯ]